



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 144269

TO: Suryaprabha Chunduru

Location: REM-2C29/2C18

Art Unit: 1637

Monday, February 14, 2005

Case Serial Number: 10/790430

From: Deirdre Arnold

Location: Biotech-Chem Library

REM 1A64

Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

RUSH

Please feel free to contact me if you have any questions or would like to amend the search.

Thank you for using STIC services.

Regards,
Deirdre Arnold

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 363.533 Seconds
(without alignments)
5324.730 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183

Sequence: 1 gacgttattatgagatgagtg...aaggaagcgcgagatgacg 1183

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database: Issued Patente NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1183	100.0	US-09-872-051-8	Sequence 8, Appl
2	202	17.1	US-09-441-340-27	Sequence 27, Appl
3	168	14.2	US-09-068-101-5	Sequence 5, Appl
4	168	14.2	US-09-970-921-5	Sequence 5, Appl
5	147	12.4	US-09-377-4668-15	Sequence 15, Appl
6	147	12.4	US-09-377-4668-13	Sequence 13, Appl
7	147	12.4	US-09-186-002-16	Sequence 16, Appl
8	147	12.4	US-09-186-002-14	Sequence 14, Appl
9	147	12.4	US-09-186-002-15	Sequence 15, Appl
10	147	12.4	US-09-186-002-13	Sequence 13, Appl
11	140	11.8	US-09-441-340-27	Sequence 27, Appl
12	140	11.8	US-09-441-340-23	Sequence 23, Appl
13	137	11.6	US-09-098-2198-5	Sequence 5, Appl
14	137	11.6	US-10-164-204-5	Sequence 5, Appl
15	137	11.6	US-09-923-109-5	Sequence 5, Appl
16	137	11.6	US-09-098-2198-6	Sequence 6, Appl
17	137	11.6	US-10-164-204-6	Sequence 6, Appl
18	137	11.6	US-09-923-109-6	Sequence 6, Appl
19	137	11.6	US-09-052-092-19	Sequence 19, Appl
20	136	11.5	US-09-068-101-6	Sequence 6, Appl
21	134	11.3	US-09-970-921-6	Sequence 6, Appl
22	134	11.3	US-09-011-151-8	Sequence 8, Appl
23	134	11.3	US-09-011-151-9	Sequence 9, Appl
24	134	11.3	US-09-794-384A-15	Sequence 15, Appl
25	134	11.3	US-09-794-384A-16	Sequence 16, Appl
26	134	11.3	US-08-064-121-3	Sequence 3, Appl
27	134	11.3	US-08-064-121-3	Sequence 3, Appl

28	134	11.3	1287	1	US-08-478-015-3	Sequence 3, Appl
29	134	11.3	1287	3	US-08-475-975-3	Sequence 3, Appl
30	134	11.3	1287	3	US-09-084-889-3	Sequence 3, Appl
31	134	11.3	1287	4	US-09-794-384A-8	Sequence 8, Appl
32	134	11.3	2345	3	US-09-026-673-1	Sequence 1, Appl
33	134	11.3	2345	3	US-09-512-650-1	Sequence 1, Appl
34	134	11.3	2345	3	US-09-480-142-1	Sequence 1, Appl
35	134	11.3	2345	4	US-09-573-555-2	Sequence 2, Appl
36	134	11.3	3544	2	US-08-485-139-3	Sequence 3, Appl
37	134	11.3	3544	2	US-08-485-139-3	Sequence 3, Appl
38	134	11.3	3544	3	US-08-750-357-3	Sequence 3, Appl
39	134	11.3	3544	3	US-08-750-357-3	Sequence 3, Appl
40	134	11.3	3658	3	US-08-894-440-3	Sequence 3, Appl
41	134	11.3	3658	3	US-08-894-440-3	Sequence 3, Appl
42	134	11.3	3658	3	US-09-458-093-3	Sequence 3, Appl
43	134	11.3	3658	3	US-09-458-093-3	Sequence 3, Appl
44	134	11.3	4032	3	US-09-068-101-5	Sequence 5, Appl
45	134	11.3	4032	4	US-09-970-921-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-872-051-8
Sequence 8, Application US/09872051
Patent No. 6823400
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
TITLE OF INVENTION: Corn Event PV-ZMGT32(hk603) and Composition and Methods
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8
Query Match
Best Local Similarity 100.0%; Score 1183; DB 4; Length 1183;
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGTCCGCAATTATACATTAAATACGC 60
DB 1 GACGTTATTATGAGATGGGTTTATGATTAGTCCGCAATTATACATTAAATACGC 60
QY 61 GATGAGAAACAAATATAGCGCGCAATAGATAATATGCGCGCGGTATCTAT 120
DB 61 GATGAGAAACAAATATAGCGCGCAATAGATAATATGCGCGCGGTATCTAT 120
QY 121 GTTACTAGATCGGGGATATCCCGGGAATTGGTACCAAGCTTTTAAATAGAGAAA 180
DB 121 GTTACTAGATCGGGGATATCCCGGGAATTGGTACCAAGCTTTTAAATAGAGAAA 180

```
Db 121 GTTACTAGATCGGGGATATCCCGGGGAATTCGGTACCAAGCTTTTATATAGTAGAAAA 180
QY 181 GAGTAATTCACCTTTGGGGCAACCTTTTATACCGATTTTATCTTATATACCACTTTTA 240
Db 181 GAGTAATTCACCTTTGGGGCAACCTTTTATACCGATTTTATCTTATATACCACTTTTA 240
QY 241 ACTGATGTTTCACTTTTGGACCAAGGTAATCTTACCTTTGTTTATTTTGAATATCCGA 300
Db 241 ACTGATGTTTCACTTTTGGACCAAGGTAATCTTACCTTTGTTTATTTTGAATATCCGA 300
QY 301 CTCTCTTCAAGCATATGATGATGACTCGAGTATGCTTTGTTAAGCGGCGCCCTTAGGA 360
Db 301 CTCTCTTCAAGCATATGATGATGACTCGAGTATGCTTTGTTAAGCGGCGCCCTTAGGA 360
QY 361 TATCAAGCTTGTATCCAGCGACACCTTCCACTAGTGTATAGTATGATGATGATGATC 420
Db 361 TATCAAGCTTGTATCCAGCGACACCTTCCACTAGTGTATAGTATGATGATGATGATC 420
QY 421 TCTTCTGAAACATTAACAGACTAGTATTTATTTGATCATTTGATCATTTGATCATTTG 480
Db 421 TCTTCTGAAACATTAACAGACTAGTATTTATTTGATCATTTGATCATTTGATCATTTG 480
QY 481 AGCGGTTTCACTTTTATTTTACAGAGCTTTTATTTTGAAGGTGCAATCCATATGCGGC 540
Db 481 AGCGGTTTCACTTTTATTTTACAGAGCTTTTATTTTGAAGGTGCAATCCATATGCGGC 540
QY 541 ATAGGTTTCACTTGGCGTATTAACCTTAACCTTAACCTTATAGCAATGCGTGTAT 600
Db 541 ATAGGTTTCACTTGGCGTATTAACCTTAACCTTATAGCAATGCGTGTAT 600
QY 601 GCGGCATCTCTTCCGCTACACAGACCTTTTACCAATCTTCTGCTGTTGCAACCCACT 660
Db 601 GCGGCATCTCTTCCGCTACACAGACCTTTTACCAATCTTCTGCTGTTGCAACCCACT 660
QY 661 GTACGAATAGCATCTACTGCTGTCTGCTGCTGCTTATTTTATTTATTAAGTAAAAACC 720
Db 661 GTACGAATAGCATCTACTGCTGTCTGCTGCTGCTTATTTTATTTATTAAGTAAAAACC 720
QY 721 ATTAATATGACAAACACACCCCTTCACTACCGGTCGAGGAGCGCCGAAGATGGGG 780
Db 721 ATTAATATGACAAACACACCCCTTCACTACCGGTCGAGGAGCGCCGAAGATGGGG 780
QY 781 TTCAACACGCTCGGACACAGATGACAGACCTTCCAGAGCAATCTGAGGCGCGGACC 840
Db 781 TTCAACACGCTCGGACACAGATGACAGACCTTCCAGAGCAATCTGAGGCGCGGACC 840
QY 841 GACGACGTAGGACAGGGGTGGCCATTAACAGACGTGGCGGATCCAACTTGTTCCTTT 900
Db 841 GACGACGTAGGACAGGGGTGGCCATTAACAGACGTGGCGGATCCAACTTGTTCCTTT 900
QY 901 TCTCTGTCTTCACTTGGCGGCGGCACTGTGCTAGACCCAGGGAATGCTGTGTGAAGAGA 960
Db 901 TCTCTGTCTTCACTTGGCGGCGGCACTGTGCTAGACCCAGGGAATGCTGTGTGAAGAGA 960
QY 961 GGTGCGGGGGCCCGATTTTATAGCTTGGCGAGAGAGCTTGGCCGAACCGATCCAGA 1020
Db 961 GGTGCGGGGGCCCGATTTTATAGCTTGGCGAGAGAGAGCTTGGCCGAACCGATCCAGA 1020
QY 1021 GCTCTGCGCAATATCAAGAAACCAAGTGGGCGCTGAGCCCAACCGCCAGAGAGC 1080
Db 1021 GCTCTGCGCAATATCAAGAAACCAAGTGGGCGCTGAGCCCAACCGCCAGAGAGC 1080
QY 1081 GGGGCTTGTTCGAGACCGGTAGCTGGGGAAGGGGACGACCCGCTAGGGGGGCCCATGCTC 1140
Db 1081 GGGGCTTGTTCGAGACCGGTAGCTGGGGAAGGGGACGACCCGCTAGGGGGGCCCATGCTC 1140
QY 1141 CACGCGCCAGAGAGAAAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
Db 1141 CACGCGCCAGAGAGAAAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
```

RESULT 2
US-09-441-340-27/c

```
Sequence 27, Application US/09441340
: Patent No. 6448476
: GENERAL INFORMATION:
: APPLICANT: Barry, Gerard F.
: TITLE OF INVENTION: Phosphonate Metabolizing Plants
: FILE REFERENCE: 38-21(15303)
: CURRENT APPLICATION NUMBER: US/09/441,340
: EARLIER FILING DATE: 1999-11-16
: EARLIER APPLICATION NUMBER: 60/108,763
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 27
: LENGTH: 2378
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: expression
: OTHER INFORMATION: cassette comprising a plant promoter linked to an
: OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
: OTHER INFORMATION: transferase, and a termination sequence
: FEATURE:
: NAME/KEY: Promoter
: LOCATION: (28)..(965)
: FEATURE:
: NAME/KEY: Intron
: LOCATION: (966)..(1423)
: FEATURE:
: NAME/KEY: transit peptide
: LOCATION: (1440)..(1667)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1668)..(2099)
: FEATURE:
: NAME/KEY: terminator
: LOCATION: (2114)..(2369)
: US-09-441-340-27
```

Query Match 17.1%; Score 202; DB 3; Length 2378;
Best Local Similarity 100.0%; Pred. No. 2,5e-89;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 163 TTTTATATATGATGAAAGATTAATTTCACTTTGGGCGCACTTTTATATACCATATTTT 222
Db 202 TTTTATATATGATGAAAGATTAATTTCACTTTGGGCGCACTTTTATATACCATATTTT 143
QY 223 ACTTATATACCACTTTTATACGATGTTTCACTTTGACAGGTAATCTTACCTTGT 282
Db 142 ACTTATATACCACTTTTATACGATGTTTCACTTTGACAGGTAATCTTACCTTGT 83
QY 283 TATTTTGAATATCCGACTCTCTTCAACGATATGATGATGATGATGATGATGATGATG 342
Db 82 TATTTTGAATATCCGACTCTCTTCAACGATATGATGATGATGATGATGATGATGATG 23
QY 343 ACGGCGCGCGCTTAGGATATC 364
Db 22 ACGGCGCGCGCTTAGGATATC 1
```

RESULT 3
US-09-068-101-5/c
Sequence 5, Application US/09068101
: Patent No. 632960
: GENERAL INFORMATION:
: APPLICANT: PLANT GENETIC SYSTEMS N.V.
: TITLE OF INVENTION: Improved Barstar Gene
: FILE REFERENCE: 2121-139P
: CURRENT APPLICATION NUMBER: US/09/068,101
: CURRENT FILING DATE: 1998-08-26
: EARLIER APPLICATION NUMBER: EP 96202446.9
: EARLIER FILING DATE: 1996-09-03
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0


```
SEQ ID NO 5
LENGTH: 4032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: other nucleic
  OTHER INFORMATION: acid, "plasmid pmv71"
FEATURE:
  NAME/KEY: misc feature
  LOCATION: (1995)..(3400)
  OTHER INFORMATION: label = PRAC1, "promoter region of rice actin gene"
  OTHER INFORMATION: - contains an intron in the leader"
FEATURE:
  NAME/KEY: misc feature
  LOCATION: (3401)..(3676)
  OTHER INFORMATION: label = barstar, "barstar DNA"
FEATURE:
  NAME/KEY: misc feature
  LOCATION: (3677)..(4003)
  OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
  OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium
  OTHER INFORMATION: T-DNA"
FEATURE:
  NAME/KEY: misc feature
  LOCATION: (3399)..(3404)
  OTHER INFORMATION: label = NcoI, "NcoI recognition site"
FEATURE:
  NAME/KEY: misc feature
  LOCATION: (4016)..(4021)
  OTHER INFORMATION: label = KpnI, "KpnI recognition site"
US-09-970-921-5

Query Match      14.2%; Score 168; DB 3; Length 4032;
Best Local Similarity 100.0%; Pred. No. 1.4e-72;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 222
    |||||
DB 2166 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 2107

QY 223 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACGAGTAATCTTACCTTGT 282
    |||||
DB 2106 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACGAGTAATCTTACCTTGT 2047

QY 283 TATTTGACTATCCGACTCTCTTTCACGATATGATGACCTCGA 330
    |||||
DB 2046 TATTTGACTATCCGACTCTCTTTCACGATATGATGACCTCGA 1999

RESULT 4
US-09-970-921-5/c
Sequence 5, Application US/09970921
Patent No. 6759575
GENERAL INFORMATION:
  APPLICANT: Frank Michiels et al.
  TITLE OF INVENTION: Improved Barstar Gene
  FILE REFERENCE: 2428-0108P
  CURRENT APPLICATION NUMBER: US/09/970,921
  NUMBER OF SEQ ID NOS: 10
  SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 4032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: other nucleic
  OTHER INFORMATION: acid, "plasmid pmv71"
  NAME/KEY: misc feature
  LOCATION: (1995)..(3400)
  OTHER INFORMATION: label = PRAC1, "promoter region of rice actin gene"
  OTHER INFORMATION: - contains an intron in the leader"
  NAME/KEY: misc feature
```

```
LOCATION: (3401)..(3676)
OTHER INFORMATION: label = barstar, "barstar DNA"
NAME/KEY: misc feature
LOCATION: (3677)..(4003)
OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium
OTHER INFORMATION: T-DNA"
NAME/KEY: misc feature
LOCATION: (3399)..(3404)
OTHER INFORMATION: label = NcoI, "NcoI recognition site"
NAME/KEY: misc feature
LOCATION: (4016)..(4021)
OTHER INFORMATION: label = KpnI, "KpnI recognition site"
US-09-970-921-5

Query Match      14.2%; Score 168; DB 4; Length 4032;
Best Local Similarity 100.0%; Pred. No. 1.4e-72;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 222
    |||||
DB 2166 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 2107

QY 223 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACGAGTAATCTTACCTTGT 282
    |||||
DB 2106 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACGAGTAATCTTACCTTGT 2047

QY 283 TATTTGACTATCCGACTCTCTTTCACGATATGATGACCTCGA 330
    |||||
DB 2046 TATTTGACTATCCGACTCTCTTTCACGATATGATGACCTCGA 1999

RESULT 5
US-09-377-466B-15
Sequence 15, Application US/09377466B
Patent No. 6501009
GENERAL INFORMATION:
  APPLICANT: Romano, Charles P.
  TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
  FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
  CURRENT APPLICATION NUMBER: US/09/377,466B
  NUMBER OF SEQ ID NOS: 43
  SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 3754
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: expression
  OTHER INFORMATION: cassette
  NAME/KEY: promoter
  LOCATION: (25)..(640)
  OTHER INFORMATION: P-CaMV.35S
  NAME/KEY: intron
  LOCATION: (669)..(1472)
  OTHER INFORMATION: I-Zm.Hsp70
  NAME/KEY: CDS
  LOCATION: (1490)..(3448)
  OTHER INFORMATION: Cry3Bb1 variant v11231
  NAME/KEY: terminator
  LOCATION: (3475)..(3730)
  OTHER INFORMATION: Agrobacterium tumefaciens nos 3' transcription
  OTHER INFORMATION: termination and polyadenylation sequence
US-09-377-466B-15

Query Match      12.4%; Score 147; DB 4; Length 3754;
Best Local Similarity 100.0%; Pred. No. 3e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGTTTATGATTTAGAGTCCCGCAATTTATACATTTAATACGC 60
    |||||
DB 3601 GACGTTATTTATGAGATGGTTTATGATTTAGAGTCCCGCAATTTATACATTTAATACGC 3660
```


Oy	61	GATGAGAAAA	CAAAATATATGCGCGCAAATAGAGATTAAATATCGCGCGGTCTCATTTAT	120
Db	3022	GATGAGAAAACAAAATATATGTCGGCGCAAACTAGAGATTAATAATTATTCGCCGCGGTCTCATTTAT	3081	
Oy	121	GTACTACTAATCGGGGATATCCCGGGGG	147	
Db	3082	GTACTAGATCGGGGATATATCCCGGGGG	3108	

RESULT 9
US-09-186-002-15

1 APPLICANT: Corbin, David R.
 2 APPLICANT: Romano, Charles P.
 3 TITLE OF INVENTION: Improved Method for Transforming Plants to Express
 4 TITLE OF INVENTION: delta-Endotoxins
 5 FILE REFERENCE: 38-21(13547) US Pat No. 5489542 09/186,002
 6 CURRENT APPLICATION NUMBER: US/05/166,002B
 7 CURRENT FILING DATE: 1998-11-04

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15

```

```

;
; LENGTH: 10252
; TYPE: DNA
; ORGANISM: Artificial Sequence
;

```

```
;
; FEATURE:
; NAME/KEY : unsure
; LOCATION: (3660) .. (3773)
;
```

```

; OTHER INFORMATION: "n" = g, a, c, or e
; FEATURE:
; NAME/KEY: unsure
;

```

US-09-186-002-15

Query Match	12.4%	Score 147	DB 4	Length 10252
Best Local Similarity	100.0%	Pred. NO.	2.9e-62	
Matches 147	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

1 GACGTTATTATGAGATGGCTTTTATGATTAGATCCCGCATTATACATTTAATACG 60

Db 2965 GACGTTATTTAGATGGGTTTTTATGATTAGATGCCGCAATTATACATTTAATATACGC 302

Db 3025 GATAGAAACAAATATAGCCGCGAATAGAGATTAATTATTCGGCGCGGTCTATCTAT 308

QY 121 GTTACTAGATCGGGGATATCCCCGGG 147
|||||
Db 3085 GTTACTAGATCGGGGATATCCCCGGG 3111

RESULT 10

US-09-186-002-13
; Sequence 13, Application US/09186002E
; Patent No. 6489542

GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Romano, Charles P.

TITLE OF INVENTION: Improved Method for Transforming Plants to Express
 TITLE OF INVENTION: delta-Endotoxins
 FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186.002

```

; CURRENT APPLICATION NUMBER: US/
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18

```

```

; SOFTWARE: Pat
; SEQ ID NO 13
; LENGTH: 10339

```

```

;
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

? FEATURE:
? NAME/KEY: unsure
? LOCATION: (3687) ..(3760)
? OTHER INFORMATION: "n"=g, a, c, or t
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (4382) ..(4434)
? OTHER INFORMATION: "n" = g, a, c, or t
US-09-186 002-13

```

Query Match	Score	DB 4	Length
12.4%	147	4	10339

```

Best Local Similarity 100.0%; Freq. NO. 2-35-02;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

DB 305 GACGTAATTATAGATGGGTTTAAATGAAGAAGCCCGCTAATAATATG.....C
DY 61 GATGAAAACAATAATATGCGGCCAACATGAGTAATATTCGCGCGGTGCATCTAT 120

Db 3112 GATGAAACAAATATATAGCGCGAACTAGATAAATATCGCGCGGTGCATCTAT 3171

Qy 121 GTTACTAGATCGGGGATAATCCCCGGG 147
|||
Db 3172 GTTACTAGATCGGGGATATCCCCGGG 3198

RESULT 11

US-09-441-340-27
; Sequence 27, Application US/09441340
; Patent No. 6448476

```

; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants

```

```
;; FILE REFERENCE: 38-21 (15303)
;; CURRENT APPLICATION NUMBER: US/09/441,340
;; CURRENT FILING DATE: 1999-11-16
```

```

; EARLIER APPLICATION NUMBER: 60/
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2378
;

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;

```

OTHER INFORMATION:	Description of Artificial Sequence:expression
OTHER INFORMATION:	cassette comprising a plant promoter linked to an
OTHER INFORMATION:	intron, a sequence encoding an AMPA acetyl

```

; OTHER INFORMATION: transferase, and a termination sequence
;
; FEATURE:
; NAME/KEY: promoter

```

```

;
; LOCATION: (28) .. (965)
; FEATURE:
; NAME/KEY: intron
;

```

```

; LOCATION: (966) .. (1423)
; FEATURE:
; NAME/KEY: transit peptide

```

```

LOCATION: (1440) .. (1667)
FEATURE:
NAME/KEY: CDS

```

```

; NAME/KEY: terminator
; LOCATION: (1668) .. (2099)
; FEATURE:
; NAME/KEY: terminator

```

LOCATION: (2114) .. (2369)
US-09-441-340-27

Query Match	11.8%	Score 140	DB 3	Length 2378
Best Local Similarity	100.0%	Pred. No	8.5e-59	
Matches 140; Conservative	0	Mismatches	0	Indels 0; Gaps 0

1 GACGTTATTATGATGGGCTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 60

Db 2239 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 2239
QY 61 GATAGAAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGTCTAT 120
Db 2299 GATAGAAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGTCTAT 2358
QY 121 GTTACTAGATCGGGGATATC 140
Db 2359 GTTACTAGATCGGGGATATC 2378

RESULT 12
US-09-441-340-31
; Sequence 31, Application US/09441340
; Patent No. 6448476
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/09/441,340
; EARLIER FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 60/108,763
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 31
; LENGTH: 2436
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: monocot
; OTHER INFORMATION: expression cassette comprising plant operable
; OTHER INFORMATION: promoter linked to an intron, a sequence coding
; OTHER INFORMATION: for an AMPA acetyltransferase, and a termination
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (26)..(640)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (670)..(1473)
; FEATURE:
; NAME/KEY: transit peptide
; LOCATION: (1498)..(1725)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1726)..(2157)
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (2172)..(2427)
; US-09-441-340-31

Query Match 11.8%; Score 140; DB 3; Length 2436;
Best Local Similarity 100.0%; Pred. No. 8.5e-57;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 60
Db 2297 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 2356
QY 61 GATAGAAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGTCTAT 120
Db 2357 GATAGAAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGTCTAT 2416
QY 121 GTTACTAGATCGGGGATATC 140
Db 2417 GTTACTAGATCGGGGATATC 2436

RESULT 13
US-09-441-340-23
; Sequence 23, Application US/09441340
; Patent No. 6448476

; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/09/441,340
; EARLIER FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 60/108,763
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression
; OTHER INFORMATION: cassette comprising plant operable promoter linked
; OTHER INFORMATION: to a coding sequence encoding an AMPA
; OTHER INFORMATION: acetyltransferase linked to a transcription
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (33)..(605)
; FEATURE:
; NAME/KEY: transit peptide
; LOCATION: (627)..(892)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (893)..(1324)
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (1350)..(1605)
; US-09-441-340-23

Query Match 11.6%; Score 137; DB 3; Length 1630;
Best Local Similarity 100.0%; Pred. No. 2.6e-57;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 60
Db 1476 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 1535
QY 61 GATAGAAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGTCTAT 120
Db 1536 GATAGAAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGTCTAT 1595
QY 121 GTTACTAGATCGGGAT 137
Db 1596 GTTACTAGATCGGGAT 1612

RESULT 14
US-09-098-2198-5
; Sequence 5, Application US/090982198
; Patent No. 6441277
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard
; APPLICANT: Cheikh, No. 6441277dine
; APPLICANT: Kishore, Ganesh
; TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/098,219B
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/049,995
 FILING DATE: 17-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Kammerer, Patricia A.
 REGISTRATION NUMBER: 29,775
 REFERENCE/DOCKET NUMBER: MOBT-086
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 713-787-1400
 TELEFAX: 713-787-1440
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10846 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-098-219B-5

Query Match 11.6%; Score 137; DB 3; Length 10846;
 Best Local Similarity 100.0%; Pred. No. 2.4e-57;
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTAGATGGGTTTATGATTAGATCCCGCAATTATATATACGC 60
 DB 4489 GACGTTATTTAGATGGGTTTATGATTAGATCCCGCAATTATATACGC 4548
 QY 61 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120
 DB 4549 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 4608
 QY 121 GTTACTAGATCGGGAT 137
 DB 4609 GTTACTAGATCGGGAT 4625

RESULT 15
 US-10-164-204-5
 Sequence 5, Application US/10164204
 Patent No. 6663906
 GENERAL INFORMATION:
 APPLICANT: Cheikh, Gerard
 APPLICANT: Kishore, Ganesh
 TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate Aldolase in Transgenic Pl
 FILE REFERENCE: 11899, 0086, DUS02 (MOBT:086-2)
 CURRENT APPLICATION NUMBER: US/10/164,204
 CURRENT FILING DATE: 2002-06-06
 PRIOR APPLICATION NUMBER: 09/098,219
 PRIOR FILING DATE: 1998-06-16
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 5
 LENGTH: 10846
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: P-FMV/CTPI/Ida/NOS3,
 US-10-164-204-5

Query Match 11.6%; Score 137; DB 4; Length 10846;
 Best Local Similarity 100.0%; Pred. No. 2.4e-57;
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTAGATGGGTTTATGATTAGATCCCGCAATTATATATACGC 60
 DB 4489 GACGTTATTTAGATGGGTTTATGATTAGATCCCGCAATTATATATACGC 4548
 QY 61 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120

DB 4549 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 4608
 QY 121 GTTACTAGATCGGGAT 137
 DB 4609 GTTACTAGATCGGGAT 4625

Search completed: February 10, 2005, 08:49:47
 Job time : 365.533 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 121.68 Seconds
(without alignments)
5630.828 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18

Sequence: 1 tgcctctcgcctccttc 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gest1.*
9: gb_gest2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	354	2	BF450993 uz73a01.y
2	18	100.0	474	1	AA103332 mo24d04.x
3	18	100.0	517	1	BM507605 ih42b06.y
4	18	100.0	518	2	BM626708 uui0e06.y
5	18	100.0	522	7	CF916734 B0999R07-
6	18	100.0	529	6	CA557138 K0222G09-
7	18	100.0	534	2	BE632036 uui0e06.x
8	18	100.0	540	4	B1966323 id50h12.x
9	18	100.0	544	5	BU605257 mah89d07-
10	18	100.0	557	7	CF171411 B0842B01-
11	18	100.0	576	2	AM559055 L0303E07-
12	18	100.0	598	4	BM219770 C0929G10-
13	18	100.0	603	7	CF739008 UI-M-HD0-
14	18	100.0	609	7	CF916509 B0996C03-
15	18	100.0	605	4	BM237875 K0508E08-
16	18	100.0	610	6	CA888455 B0145C12-
17	18	100.0	617	2	BM601036 BM601036
18	18	100.0	617	4	BM507714 ih42b06.x
19	18	100.0	629	4	B1650935 603297263
20	18	100.0	629	5	BU702841 UI-M-FC0-
21	18	100.0	630	4	BM076391 H3159C09-
22	18	100.0	630	4	BM233685 K0340B05-
23	18	100.0	668	7	CNS31891 UI-M-HD0-
24	18	100.0	709	5	BQ770022 UI-M-F10-

25	18	100.0	742	4	BI853956 603381039
26	18	100.0	812	4	BI872911 603398262
27	18	100.0	907	2	BF179971 601806316
28	18	100.0	997	5	BQ713665 AGENCOURT
29	18	100.0	997	5	AK036132 Mus muscu
30	18	100.0	997	5	AK036132 Mus muscu
31	17	94.4	604	4	BM236573
32	17	94.4	655	8	AZ572271 304PvB11
33	16	88.9	182	5	BQ321075 CM0-CT030
34	16	88.9	182	5	BQ321075 CM0-CT030
35	16	88.9	297	8	AQ421868 RPCI-11-1
36	16	88.9	433	8	AQ298483 HS 2229 B
37	16	88.9	503	2	BF289471 EST454062
38	16	88.9	537	9	CL409705 RPCI4_41
39	16	88.9	646	9	BX197672 Danio rer
40	16	88.9	653	9	BX122080 Danio rer
41	16	88.9	672	8	AZ835346 2M0129H05
42	16	88.9	761	9	BX142562 Danio rer
43	16	88.9	782	5	CA384083 BX913369
44	16	88.9	796	9	CA384083 664470 NC
45	16	88.9	811	8	B2175244 Tetradon

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BF450993 354 bp mRNA linear EST 29-DEC-2000
uz73a01.y1 NCI_CGAP_Lu29 Mus musculus CDNA clone IMAGE:3674664 5',
mRNA sequence.
BF450993 GI:11517162
EST.
Mus musculus (house mouse)
Mus musculus
Bukavota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 354)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: uz73a01.x1
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.llnl.gov/image/html/iresources.shtml
MGI:1435432
Seq primer: -40RP from Gibco
High quality sequence stop: 340.
Location/Qualifiers
1. 354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6 J"
/db_xref="taxon:10090"
/clone="IMAGE:3674664"
/tissue_type="spontaneous tumor, metastatic to mammary."
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 354;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
 |||||
 40 TGCTGTTCTGCTGACTTT 57

RESULT 2
 AA103332 474 bp mRNA linear EST 29-OCT-1996
 LOCUS mo2404.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus
 DEFINITION cDNA clone IMAGE:554503 5', mRNA sequence.

ACCESSION AA103332
 VERSION AA103332.1 GI:1649616
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Mammalia, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 474)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watscn.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:335295
 Seq primer: -28M13 rev1 from Amersham
 High quality sequence stop: 430.
 Location/Qualifiers

FEATURES
 source
 1..474
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:554503"
 /tissue_type="embryo"
 /dev_stage="13.5dpc embryos"
 /lab_host="DH10B"
 /clone_idb="Life Tech mouse embryo 13 5dpc 10666014"
 /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site: 1; Salt; Site 2: NCI; Cloned unidirectionally. Primer: Oligo dt. 13.5dpc embryos. pCMV-SPORT2 vector."

ORIGIN
 Query Match 100.0%; Score 18; DB 1; Length 474;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
 |||||
 337 TGCTGTTCTGCTGACTTT 354

RESULT 3
 BM507605 517 bp mRNA linear EST 12-MAR-2002
 LOCUS h42506.y1 Melton Mouse E16 5 Pancreas Library 2 M1652 Mus musculus
 DEFINITION cDNA clone IMAGE:5679706 5', mRNA sequence.

ACCESSION BM507605

VERSION BM507605.1 GI:18678748
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 517)
 Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K., Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Bilestein, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, K., Tsagaris, R., Williams, T., Jackson, Y., and Bowers, Y.
 The Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu)
 MGI:1957584 This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov
 Seq primer: -40RP from Gibco
 High quality sequence stop: 432.
 Location/Qualifiers

FEATURES
 source
 1..517
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129"
 /db_xref="taxon:10090"
 /clone="IMAGE:5679706"
 /sex="Both"
 /tissue_type="Total pancreas"
 /dev_stage="Embryonic day 16.5"
 /lab_host="TOP10"
 /clone_idb="Melton Mouse E16 5 Pancreas Library 2 M1652"
 /note="Organ: Pancreas; Vector: pBluescript II SK; Site: 1; NCI; Site 2: Salt; Library constructed using Superscript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.06Kb. Primary library, unamplified."

ORIGIN
 Query Match 100.0%; Score 18; DB 4; Length 517;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
 |||||
 430 TGCTGTTCTGCTGACTTT 447

RESULT 4
 BE626708 518 bp mRNA linear EST 24-AUG-2000
 LOCUS u110e06.y2 Soares mammary_gland NMIMG Mus musculus cDNA clone
 DEFINITION IMAGE:3371554 5' similar to SW:KRU1_DROME P08155 KRUEPPEL HOMOLOGOUS PROTEIN; , mRNA sequence.

ACCESSION BE626708
 VERSION BE626708.1 GI:9907124
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 518)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapdb-@email.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:1081158
Seq primer: -40RP from Gibco
High quality sequence stop: 456.
Location/Qualifiers
1. 518
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3371554"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo (dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified p7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 100.0%; Score 18; DB 2; Length 518;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 493 TGCTGTTCTGCTGACTTT 510

RESULT 5
CF916734 522 bp mRNA linear EST 05-NOV-2003
LOCUS B0999F07-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus
DEFINITION musculus cDNA clone NIA:B0999F07 IMAGE:30481794 5', mRNA sequence.
CF916734
ACCESSION CF916734.1 GI:38187936
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 522)
AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from
subprogram amounts of total RNAs by a universal PCR amplification
method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0999 row: F column: 07
Seq primer: M13 Reverse
High quality sequence stop: 522
POLYA=No.
Location/Qualifiers
1. 522
/organism="Mus musculus"

FEATURES
SOURCE

/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0999F07-5"
/db_xref="taxon:10090"
/clone="NIA:B0999F07 IMAGE:30481794"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (<http://igsun.grc.nia.nih.gov/cDNA>).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 1488 unfertilized eggs.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen):
5'-pGACTGTTCTAGATCGAGCGCGCCCTTTTTTTTTTT-3').
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker L-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer SalI-4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

ORIGIN
Query Match 100.0%; Score 18; DB 7; Length 522;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 183 TGCTGTTCTGCTGACTTT 200

RESULT 6
CA557138 529 bp mRNA linear EST 19-NOV-2002
LOCUS K0222G09-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus
DEFINITION musculus cDNA clone NIA:K0222G09 IMAGE:30046832 5', mRNA sequence.
CA557138
ACCESSION CA557138.1 GI:25101573
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 529)
AUTHORS Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Stagg, C.A.,
Martin, P., Alba, K., Tanaka, T. and Ko, M.S.H.
TITLE Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library
(Long)
JOURNAL Unpublished (2001)
COMMENT Other ESTs: K0222G09-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0222 row: G column: 09
Seq primer: M13 Reverse
High quality sequence stop: 529
POLYA=No.
Location/Qualifiers
1. 529

FEATURES
SOURCE

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="NIA:K0222G09 IMAGE:30046832"
/issue_type="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001) [PMID: 11544191]). Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):
5'-GGAAGTCTTCAGATCGGAGCGCCGCTTTTCTTTTCTTTT-3',
treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 529;
Best Local Similarity, 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 211 TGCTGTTCTGCTGACTTT 228

RESULT 7
BE632036/c 534 bp mRNA linear EST 25-AUG-2000
LOCUS un10e06.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:3371554 3', mRNA sequence.
ACCESSION BE632036
VERSION BE632036.1 GI:9914724
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 534)
NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: un10e06.y2
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LIND; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1081158
High quality sequence stop: 469.
Location/Qualifiers
1..534
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

FEATURES

source
1..534
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

/clone="IMAGE:3371554"
/sex="female (lactating)"
/issue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptor (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library is normalized. Library was constructed by Bento Soares and W. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 506 TGCTGTTCTGCTGACTTT 489

RESULT 8
B1966323/c 540 bp mRNA linear EST 12-MAR-2002
LOCUS i450h12.x1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus
DEFINITION musculus cDNA clone IMAGE:5666926 3', mRNA sequence.
ACCESSION B1966323
VERSION B1966323.1 GI:16340728
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 540)
Melton, D., Brown, J., Keny, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scarce, M., Brechtel, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A., Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu)
MGI:1953252 This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 449.
Location/Qualifiers
1..540
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5666926"
/sex="Both for embryonic & newborn, male for adult and adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"
/lab_host="DH10B"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1

TITLE

JOURNAL

COMMENT

FEATURES

source

1..540
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5666926"
/sex="Both for embryonic & newborn, male for adult and adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"
/lab_host="DH10B"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1

NI-MMS1"
 /note="Vector: pSPORT1; Site 1: Not 1; Site 2: Sal I; Five
 libraries representing E10.5/12.5 pancreatic bud, E16.5
 pancreas, newborn pancreas, adult pancreas, and adult
 islets of Langerhans were separately constructed using
 Superscript Plasmid Library kit (Life Technologies). cDNA
 was made by oligo-dT priming and size-selected by column
 fractionation. Libraries were amplified once on solid
 support and plasmid DNA from each library was prepared
 and mixed in equal amounts. The mixed library DNA was
 normalized by method #4 from Bonaldo, Lennon, and Soares
 1996 Genome Research 6:791-806; 0.5 microgram
 single-stranded mixed library plasmid DNA was mixed with
 5 micrograms PCR product representing mixed library
 inserts and hybridized to an EcoT of 6. Single-stranded
 (unhybridized) plasmids were isolated by hydroxyapatite
 chromatography and used to make this library."

ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 540;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
 |||||
 Db 504 TGCTGTTCTGCTGACTTT 487

RESULT 9 544 bp mRNA linear EST 20-SEP-2002
 BU605257 mah89b07.y1 McCarrey Eddy 18 day leptotene and zygotene
 LOCUS spermatocytes Mus musculus cDNA IMAGE:6367453 5', mRNA
 DEFINITION
 accession.
 BU605257
 BU605257.1 GI:23261437

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 McCarrey, J., Eddy, M., Marra, M., Hillier, L., Clifton, S., Pape, D.,
 Martin, J., Wylie, T., Dante, M., Bowers, Y., Theising, B., Gibbons, M.,
 Rutter, E., Tsagaris, R., Ronko, I., Maguire, L., Kennedy, S.,
 Bennett, J., Waterston, R. and Wilson, R.

TITLE
 JOURNAL
 COMMENT
 Unpublished (2002)
 Contact: McCarrey/Eddy NIEHS Mouse
 NIEHS Mouse
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

FEATURES
 source
 Location/Qualifiers
 1..544
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6367453"
 /sex="male"
 /issue_type="18-day leptotene and zygotene spermatocytes"
 /lab_host="DHI08 (phage-resistant)"
 /clone_lib="McCarrey Eddy 18 day leptotene and zygotene
 spermatocytes"

/note="Organ: testis; Vector: pBluescript SK+
 (Stratagene); Site 1: EcoRI; Site 2: XhoI; cDNA oligo
 dT-primed [5'-(GA)10-ACGTGCTGCGATTTT-3'] and
 directionally cloned using 5' linkers 5'-ATTGCGACGAG-3'
 and 5'-CTCGGCG-3'. Size selection of >400bp material
 gives average insert size ranging from 1-2 kb. Library was
 mass excised (from lambda-UniZAP-XR) and resulting
 single-stranded phagmids were prepared and transformed into
 DH10B. Library constructed and donated by J. McCarrey,
 Ph.D. (Southwest Foundation for Biomedical Research, Dept.
 of Genetics); excision done by E.M. Eddy, Ph.D. (National
 Institutes of Health, National Institute of Environmental
 Health Sciences)."

ORIGIN

Query Match 100.0%; Score 18; DB 5; Length 544;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
 |||||
 Db 341 TGCTGTTCTGCTGACTTT 358

RESULT 10 557 bp mRNA linear EST 25-JUL-2003
 CP171411 B0842B01-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
 LOCUS musculus cDNA clone NIA:B0842B01 IMAGE:30471468 5', mRNA sequence.
 DEFINITION
 CP171411
 CP171411.1 GI:33280960

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 1 (bases 1 to 557)
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: B0842 row: B column: 01

Seq primer: M13 Reverse
 High quality sequence stop: 557
 POLYA=No.
 Location/Qualifiers
 1..557
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30471468"
 /dev_stage="Newborn Kidney"
 /lab_host="DHI08"
 /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
 1)"

FEATURES
 source
 Location/Qualifiers
 1..557
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30471468"
 /dev_stage="Newborn Kidney"
 /lab_host="DHI08"
 /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
 1)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
 Site 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
 In brief, double-stranded cDNAs were synthesized with an
 oligo(dT) primer (Invitrogen):
 5'-TGACTGATTCAGATCGGACGCGCCGCTTTT-3' from
 26 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-5. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 557;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGACTTT 18
DB 447 TGCTGTTCTGCTGACTTT 464

RESULT 11
AM559055/c 576 bp mRNA linear EST 31-AUG-2000
LOCUS 10303E07-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA
DEFINITION clone 10303E07 3', mRNA sequence.
ACCESSION AM559055
VERSION AM559055.1 GI:7204484
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X.,
Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagarsaja, R., Doi, R.,
Wood, W.H. III, Becker, K.G. and Ko, M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (Feb. 9, 2000) 9127-9132 (2000)

TITLE

JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: 10303 row: E column: 07
Seq primer: -21M13 Forward
High quality sequence stop: 576
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..576
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="10303E07"
/sex="Female"
/dev_stage="Newborn Ovary"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Ovary cDNA Library"
/note="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
7 Newborn Ovary. The double-stranded cDNA was synthesized
by Gibco's kit with an Oligo(dT) primer [NotI
15'-pGACTGTTCTAGATCGGCGCGCCCTTTTCTTTTCTTT-3'] from
2.5ug of total RNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by

ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Yulan Piao."

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 576;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGACTTT 18
DB 520 TGCTGTTCTGCTGACTTT 503

RESULT 12
BM219770/c 598 bp mRNA linear EST 07-JUN-2003
LOCUS BM219770
DEFINITION C0929G10-3 NIA Mouse 12.5-dpc Male Genital Ridge/Mesonephros cDNA
Library (long) Mus musculus cDNA clone NIA:C0929G10 IMAGE:30037425
3', mRNA sequence.
ACCESSION BM219770
VERSION BM219770.2 GI:31478468
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL
MEDLINE
PUBMED
COMMENT
On Dec 14, 2001 this sequence version replaced gi:1779692.
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: C0929 row: G column: 10
Seq primer: -21M13 Forward
High quality sequence stop: 598
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..598
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="NIA:C0929G10 IMAGE:30037425"
/sex="Male"
/issue_type="Male genital ridge/mesonephros"
/dev_stage="12.5-dpc"
/lab_host="DH10B"
/clone_lib="NIA Mouse 12.5-dpc Male Genital
Ridge/Mesonephros cDNA Library (long)"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is

separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 609;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 211 TGCTGTTCTGCTGACTTT 228

RESULT 15 609 bp mRNA linear EST 07-JUN-2003
BM237875 K0508E08-3 NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit+/Sca-1+)
LOCUS CDNA Library (Long) Mus musculus CDNA clone NIA:K0508E08
DEFINITION IMAGE:30064663 3', mRNA sequence.

ACCESSION BM237875 GI:31494430

VERSION BM237875.2

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE

AUTHORS

TITLE

Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)

MEDLINE 21429098

PUBMED 11544199

COMMENT On Dec 17, 2001 this sequence version replaced gi:17873145.

Other ESTs: K0508E08-5N

Contact: Dawood B. Dudekula

Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: K0508 row: E column: 08

Seq primer: -21M13 Forward

High quality sequence stop: 609

POLYA=yes

Location/Qualifiers

1. .609

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6Ncr"

/db_xref="niaEST:K0508E08-3"

/db_xref="taxon:10090"

/clone="NIA:K0508E08 IMAGE:30064663"

/tissue_type="Hematopoietic Stem Cell

(Lin-/C-Kit+/Sca-1+)"

/dev_stage="Age approx.10 weeks old"

/lab_host="DH10B"

/clone_id="NIA Mouse Hematopoietic Stem Cell

(Lin-/C-Kit+/Sca-1+) CDNA Library (Long)"

/note="Vector: pSPORT (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is

a long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). [PMID: 11544199]. Total RNAs were

obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):

5'-pGACTACTTCTAGATCGGAGCGCGCCGCTTTTCTTTT-3' from 4.8 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.7 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 504 TGCTGTTCTGCTGACTTT 487

Search completed: February 10, 2005, 17:01:59
Job time : 124.68 secs

Db 181 AGCCGCGTGCATTCGGCCAGACTCTCTCTCGGCATGAGCCGATCTTTCTCTG 240
QY 241 GCATTTCCAAACCCAGAGAGTGGTCCCTGGTGGCTCCGGCCAGCAAGCTTGTAG 300
Db 241 GCATTTCCAAACCCAGAGAGTGGTCCCTGGTGGCTCCGGCCAGCAAGCTTGTAG 300
QY 301 CGGCCCAACGGGTGATCCAGCTTGATCCCTAGGGCGCGCGCTTAAACAAGCTTACTC 360
Db 301 CGGCCCAACGGGTGATCCAGCTTGATCCCTAGGGCGCGCGCTTAAACAAGCTTACTC 360
QY 361 GAGGTCAATCATATGCTTGAGAGAGAGTGGGATAGTCCAAAATTAACAAGGTAGA 420
Db 361 GAGGTCAATCATATGCTTGAGAGAGAGTGGGATAGTCCAAAATTAACAAGGTAGA 420
QY 421 TTACCCGTCAAAGGTGAAAACATCATGTTAAAGGTAAAGTAAATATCGTAAATA 480
Db 421 TTACCCGTCAAAGGTGAAAACATCATGTTAAAGGTAAAGTAAATATCGTAAATA 480
QY 481 AAGGTGCGCCAAAGTGAA 498
Db 481 AAGGTGCGCCAAAGTGAA 498

RESULT 2
AX342369/c 1183 bp DNA linear PAT 12-JAN-2002
LOCUS AX342369
DEFINITION Sequence 8 from Patent EP1167531.
ACCESSION AX342369
VERSION AX342369.1 GI:18151812
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 Behr, C.F., Hironaka, C., Heck, G.R. and You, J.
AUTHORS Corn transformant pv-zmg132 (nk603) and compositions and methods
TITLE for detection thereof
JOURNAL Patent: EP 1167531-A 8 02-JAN-2002;
Monsanto Technology LLC (US)
FEATURES
source Location/Qualifiers
1. .1183
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="1-164 Agrobacterium tumefaciens nos 3' terminator
165-381 construct vector DNA 382-686 Zea mays genomic DNA"
genes, tpst1 and ipoa 687-1183 Zea mays genomic DNA"
ORIGIN
Query Match 23.7%; Score 118; DB 6; Length 1183;
Best Local Similarity 100.0%; Pred. No. 1.8e-53;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 308 CGCGTGTACCAAGCTTGATTCCTTAGGGCGCGCGCTTAAACAAGCTTACTCGAGTCA 367
Db 381 CGCGTGTACCAAGCTTGATTCCTTAGGGCGCGCGCTTAAACAAGCTTACTCGAGTCA 322
QY 368 TTCTATATGCTTGAGAGAGAGTGGGATAGTCCAAAATTAACAAGGTAGAATTACC 425
Db 321 TTCTATATGCTTGAGAGAGAGTGGGATAGTCCAAAATTAACAAGGTAGAATTACC 264
RESULT 3
AR229555 2378 bp DNA linear PAT 20-DEC-2002
LOCUS AR229555
DEFINITION Sequence 27 from Patent US 6448476.
ACCESSION AR229555
VERSION AR229555.1 GI:27269171
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 2378)
AUTHORS Barry, G.F.
TITLE Plants and plant cells transformation to express an
JOURNAL AMPA-N-acetyltransferase
FEATURES Patent: US 6448476-A 27 10-SEP-2002;
source Location/Qualifiers
1. .2378
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 20.3%; Score 101; DB 6; Length 2378;
Best Local Similarity 100.0%; Pred. No. 4.7e-44;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 325 GATATCCCTAGGGCGCGCGCTTAAACAAGCTTACTCGAGTCAATCATATGCTTGAGAG 384
Db 1 GATATCCCTAGGGCGCGCGCTTAAACAAGCTTACTCGAGTCAATCATATGCTTGAGAG 60
QY 385 AGAGTCGGATAGTCCAAAATTAACAAGGTAGAATTACC 425
Db 61 AGAGTCGGATAGTCCAAAATTAACAAGGTAGAATTACC 101

RESULT 4
S44221 1266 bp DNA linear PLN 18-MAR-1998
LOCUS S44221
DEFINITION Act1-actin 1 (5' region, promoter) [rice, Genomic, 1266 nt].
ACCESSION S44221
VERSION S44221.1 GI:256131
KEYWORDS
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE 1 Eukaryota: Viridiplantae; Streptophyta: Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
TITLE 1 (bases 1 to 1266)
McElroy, D., Zhang, W., Cao, J. and Wu, R.
JOURNAL Isolation of an efficient actin promoter for use in rice
MEDLINE Plant Cell 2 (2), 163-171 (1990)
PUBMED 92404732
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbs 114121] from the original journal article.
This sequence comes from Fig. 2.
FEATURES
source Location/Qualifiers
1. .1266
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
1246. .1266
/gene="Act1"
/note="actin 1"
1246. .1266
/gene="Act1"
/note="actin 1"
/codon_start=1
/protein_id="AA013837.1"
/db_xref="GI:4261543"
/translation="MADAEDI"
ORIGIN
Query Match 14.1%; Score 70; DB 8; Length 1266;
Best Local Similarity 100.0%; Pred. No. 6.7e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGTCCGAGATAGTCCAAAATTAACAAGG 415
Db 10 TACTCGAGTCATTCATATGCTTGAGAGAGTCCGAGATAGTCCAAAATTAACAAGG 69
QY 416 TAAGATTACC 425
|||||

Db 70 TAAGATTACC 79

RESULT 5
LOCUS AX044144 150115
DEFINITION Sequence 7 from patent US 5641876.
ACCESSION 150115
VERSION 150115.1 GI:2472335
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1404)
AUTHORS McElroy,D. and Wu,R.
TITLE Rice actin gene and promoter
JOURNAL Patent: US 5641876-A 7 24-JUN-1997;
FEATURES
source
1. .1404
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 14.1%; Score 70; DB 6; Length 1404;
Best Local Similarity 100.0%; Pred. No. 6.7e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
|||||
9 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 68
|||||

Db 416 TAAGATTACC 425
|||||
69 TAAGATTACC 78
|||||

RESULT 6
LOCUS AX044094 150113
DEFINITION Sequence 49 from Patent W00066748.
ACCESSION AX044094
VERSION AX044094.1 GI:11342972
KEYWORDS
SOURCE Oryza sp.
ORGANISM Oryza sp.
REFERENCE 1
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and
TITLE Pickering,I.L.A.P.
JOURNAL Herbicide resistant plants
Patent: WO 0066748-A 49 09-NOV-2000;
FEATURES
source
1. .1501
Location/Qualifiers
/organism="Oryza sp."
/mol_type="unassigned DNA"
/db_xref="taxon:52841"

ORIGIN
Query Match 14.1%; Score 70; DB 6; Length 1501;
Best Local Similarity 100.0%; Pred. No. 6.7e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
|||||
750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 809
|||||

QY 416 TAAGATTACC 425
|||||
810 TAAGATTACC 819
|||||

RESULT 7
LOCUS AX044144 150113
DEFINITION Sequence 44 from Patent W00066747.
ACCESSION AX044144
VERSION AX044144.1 GI:11343022
KEYWORDS
SOURCE Oryza sp.
ORGANISM Oryza sp.
REFERENCE 1
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and
TITLE Pickering,I.L.A.P.
JOURNAL Herbicide resistant plants
Patent: WO 0066747-A 44 09-NOV-2000;
FEATURES
source
1. .1501
Location/Qualifiers
/organism="Oryza sp."
/mol_type="unassigned DNA"
/db_xref="taxon:52841"

ORIGIN
Query Match 14.1%; Score 70; DB 6; Length 1501;
Best Local Similarity 100.0%; Pred. No. 6.7e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
|||||
750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 809
|||||

Db 416 TAAGATTACC 425
|||||
810 TAAGATTACC 819
|||||

RESULT 8
LOCUS 150113 2199 bp DNA
DEFINITION Sequence 5 from patent US 5641876.
ACCESSION 150113
VERSION 150113.1 GI:2472333
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1
AUTHORS McElroy,D. and Wu,R.
TITLE Rice actin gene and promoter
JOURNAL Patent: US 5641876-A 5 24-JUN-1997;
FEATURES
source
1. .2199
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 14.1%; Score 70; DB 6; Length 2199;
Best Local Similarity 100.0%; Pred. No. 6.8e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
|||||
806 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 865
|||||

QY 416 TAAGATTACC 425
|||||
866 TAAGATTACC 875
|||||

RESULT 9

AY452735 4895 bp DNA circular SYN 19-DEC-2003
 LOCUS Reporter vector pACTXN, complete sequence.
 DEFINITION AY452735
 ACCESSION AY452735
 VERSION AY452735.1 GI:39636979
 KEYWORDS
 SOURCE Reporter vector pACTXN
 ORGANISM Reporter vector pACTXN
 other sequences; artificial sequences; vectors.
 REFERENCE
 1 (bases 1 to 4895)
 Vickers, C.E., Xue, G.P. and Gresshoff, P.M.
 A synthetic xylanase as a novel reporter in plants
 JOURNAL Plant Cell Rep. 22 (2), 135-140 (2003)
 MEDLINE 22867549
 PUBMED 12845475
 REFERENCE
 2 (bases 1 to 4895)
 Vickers, C.E.
 Direct Submission
 Submitted (29-OCT-2003) ARC Centre for Integrative Legume Research,
 The University of Queensland, Room 213, John Hines Building (69),
 St. Lucia, QLD 4072, Australia
 FEATURES
 source Location/Qualifiers
 1..4895
 /organism="Reporter vector pACTXN"
 /mol_type="other DNA"
 misc_feature /db_xref="taxon:255329"
 267..2274
 /note="sequence from cloning vector pBR322"
 468..1328
 /note="ampicillin resistance protein; bla"
 /codon_start=1
 /product="beta-lactamase"
 /protein_id="AAR29085.1"
 /db_xref="GI:39636980"
 /translation="MSIQHRRVALIPPPAFCUPVPAHPETLVKVAADOLGARVGY
 IEIDLNSGKILLESFPEREPMSFTKVLICGAVLSRIDAQOEQLRIRHYSQNDLVE
 YSPVTEKHLTDGMVRELCSAATMSDNTANLLITIGGKELTAFLEHMGDVTRL
 DRWPEPELNEAI PNDERDTTPVAMATTLRLTGELTLASROQLDMMEDAKVAGPL
 LRSALPGWFIADKSGAGERSGIIALGPDKPSRIIVITTSQATMDERNRQIA
 EIGASLIKHW"
 misc_feature 2345..2367
 /note="T7 RNA polymerase promoter"
 2426..3338
 /note="from rice actin gene; Act1"
 3339..3792
 /note="intron 1 from rice actin gene; Act1"
 3800..4513
 /gene="sxynA"
 /cDS 3800..4513
 /gene="sxynA"
 /note="synthetic xylanase"
 /codon_start=1
 /product="Xylanase"
 /protein_id="AAR29086.1"
 /db_xref="GI:39636981"
 /translation="MASNGKKFTVGNQGNQKGVNDGSEYIEIDNTGNSMTLGSG
 ATPAEENAAVNRGNPLARGLDPSOKADYDYGADVYATKOTASAGNRLCY
 YGWFONGNLGCVPLVEYIIEEDWDVDPDAGKVTIDGAGYKIRQMDHNSPTINGGS
 ETRKQYPSVRQOKRTSGHITVSDHFKERAKGSMGIGNLYEVALNAEGRQSGVADVTL
 LDVYTPKSGSPATSAAPR"
 4535..4872
 /note="from rDcs gene"
 4895
 /note="SP6 RNA polymerase transcription initiation site"
 ORIGIN
 Query Match 14.1%; Score 70; DB 12; Length 4895;
 Best Local Similarity 100.0%; Pred. No. 6.9e-27;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 TACTCGAGGTCATTCATATGCTTGAGAGAGAGAGTCGGATAGTCCAAAATAAACAAGG 415
 DB 2423 TACTCGAGGTCATTCATATGCTTGAGAGAGAGAGTCGGATAGTCCAAAATAAACAAGG 2482

QY 416 TAAGATTACC 425
 DB 2483 TAAGATTACC 2492
 RESULT 10
 150112 5643 bp DNA 11near PAT 07-OCT-1997
 LOCUS Sequence 4 from patent US 5641876.
 DEFINITION 150112
 ACCESSION 150112
 VERSION 150112.1 GI:2472332
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 5643)
 McElroy, D. and Wu, R.
 Rice actin gene and promoter
 Patent: US 5641876-A 4 24-JUN-1997;
 FEATURES
 source Location/Qualifiers
 1..5643
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 14.1%; Score 70; DB 6; Length 5643;
 Best Local Similarity 100.0%; Pred. No. 6.9e-27;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 TACTCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAAATAAACAAGG 415
 DB 809 TACTCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAAATAAACAAGG 868
 QY 416 TAAGATTACC 425
 DB 2483 TAAGATTACC 2492
 RESULT 11
 AX384394 9359 bp DNA 1linear PAT 19-MAR-2002
 LOCUS Sequence 3 from Patent W00214524.
 DEFINITION AX384394
 ACCESSION AX384394
 VERSION AX384394.1 GI:19577678
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.
 REFERENCE
 1
 van der Valk, P., van Dun, C.M., Smeekens, S.C. and Proveniers, M.C.
 Inhibition of generative propagation in genetically modified
 herbicide resistant grasses
 Patent: WO 0214524-A 3 21-FEB-2002;
 Advanta Seeds B.V. (NL)
 JOURNAL Location/Qualifiers
 1..9359
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32650"
 839..1699
 /note="unnamed protein product; Beta-lactamase gene
 (AmpR)"
 /codon_start=1
 /transl_table=1
 /protein_id="CAD28571.1"
 /db_xref="GI:19577679"
 /translation="MSIQHRRVALIPPPAFCUPVPAHPETLVKVAADOLGARVGY
 IEIDLNSGKILLESFPEREPMSFTKVLICGAVLSRIDAQOEQLRIRHYSQNDLVE
 YSPVTEKHLTDGMVRELCSAATMSDNTANLLITIGGKELTAFLEHMGDVTRL
 DRWPEPELNEAI PNDERDTTPVAMATTLRLTGELTLASROQLDMMEDAKVAGPL
 LRSALPGWFIADKSGAGERSGIIALGPDKPSRIIVITTSQATMDERNRQIA
 EIGASLIKHW"

promoter 2941..4920
/note="Ubi-promoter from maize"
misc_feature 4921..6400
/note="ACHI gene from Arabidopsis thaliana"
polyA_signal 6401..6672
/note="Poly-A signal from the nopaline synthetase gene
from Agrobacterium tumefaciens"
misc_feature 7434..8084
/note="First exon-intron combination from Ubi-maize"
misc_feature 8085..9119
/note="Hygromycin resistance gene from Escherichia coli"
polyA_signal 9120..9359

Query Match 14.1%; Score 70; DB 6; Length 9359;
Best Local Similarity 100.0%; Pred. No. 7e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATATGTCGAAATTAACCAAGG 415
|||||
DB 6688 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATATGTCGAAATTAACCAAGG 6747
|||||

QY 416 TAAAGTTACC 425
|||||
DB 6748 TAAAGTTACC 6757
|||||

RESULT 12
AX473364 9359 bp DNA linear PAT 09-AUG-2002
LOCUS AX473364
DEFINITION Sequence 1 from Patent W00214486.
ACCESSION AX473364
VERSION AX473364.1 GI:22207993
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS van der Valk, P., van Dun, C.M., Smeekens, S.C. and Proveniers, M.C.
TITLE Inhibition of generative propagation in genetically modified
grasses
JOURNAL Patent: WO 0214486-A 1 21-FEB-2002;
Advanta Seeds B.V. (NL)
FEATURES
source Location/Qualifiers
1..9359
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
2941..4920
/note="Ubi-promoter from maize"
6401..6672
/note="Poly-A signal from the nopaline synthetase gene
from Agrobacterium tumefaciens"
7434..8084
/note="First exon-intron combination from Ubi-maize"
9120..9359
ORIGIN
Query Match 14.1%; Score 70; DB 6; Length 9359;
Best Local Similarity 100.0%; Pred. No. 7e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATATGTCGAAATTAACCAAGG 415
|||||
DB 6688 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATATGTCGAAATTAACCAAGG 6747
|||||

QY 416 TAAAGTTACC 425
|||||
DB 6748 TAAAGTTACC 6757
|||||

RESULT 13
AC091532

LOCUS AC091532 147344 bp DNA linear PLN 11-FEB-2003
DEFINITION Oryza sativa chromosome 3 BAC OSJNBa0078A17 genomic sequence,
complete sequence.
ACCESSION AC091532
VERSION AC091532.13 GI:28173090
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 147344)
AUTHORS Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, R.M.,
Padrosch, D.W., Tallon, L.J., Koo, H., Ziemann, V., Heiao, J., Blunt, S.,
Vanaken, S.S., Riedmuller, S.B., Utecherback, T.T., Feldlyum, T.V.,
Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.,
White, O., Salzberg, S.L. and Fraser, C.M.
Oryza sativa chromosome 3 BAC OSJNBa0078A17 genomic sequence
Unpublished
2 (bases 1 to 147344)
AUTHORS Buell, R.
REFERENCE Direct Submission
TITLE Submitted (12-NOV-2002) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 147344)
AUTHORS Buell, R.
REFERENCE Direct Submission
TITLE Submitted (03-MAY-2001) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
4 (bases 1 to 147344)
AUTHORS Buell, R.
REFERENCE Direct Submission
TITLE Submitted (31-JAN-2003) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
5 (bases 1 to 147344)
AUTHORS Buell, R.
REFERENCE Direct Submission
TITLE Submitted (11-FEB-2003) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
COMMENT On Jan 31, 2003 this sequence version replaced gi:24308660.
Address all correspondence to: rice@tigr.org

BAC clone OSJNBa0078A17 is from Oryza sativa chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (<http://www.softberry.com/>),
GENSCAN and Genscan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GenemarkHM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/genemark/>), and GeneSplicer
(Mihela Petrea and Steven Salzberg, [contact mperetia.tigr.org](http://contact.mperetia.tigr.org)),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
identified by repeatmasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).
This BAC overlaps with rice BACOSJNBa0014G15 (AC090882).
FEATURES
source Location/Qualifiers
1..147344
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"

/chromosome="3"
/map="C10769"
/clone="OSJNBa0078A17"
/note="japonica cultivar-group"
complement(890..3895)
/gene="OSJNBa0078A17.1"
/note="predicted by fgenesh"
complement(join(<890..2270,2551..2797,3430..3895))
/gene="OSJNBa0078A17.1"
complement(join(890..2270,2551..2797,3430..3895))
/gene="OSJNBa0078A17.1"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA038818.1"
/db_xref="GI:28301925"
/translation="MPSGMRALERLASMPTEPPPPPPPPAKAAAEAREPCPSAA
TTSSSSSSASAPAAAGAPRVLPADPPEGAAGLGGAPDNTWMLRQMGASPS
PAAGPGSAGAAVAVNVRVLSDRTGYVVIYFAVEEVASPEPLCDYCNWAGSMH
WSEKRYHFIIPADWDRLADVILKRTDHLHGLHNSGFHLVMLRGDGSITL
LSGRDINDWDCSALRAVAVVDFSQKSLDLRLLSVAGDVTWETRWYGLARG
CFVSTSTYAAVVALAALPYDYSRSHRVVITRSLNKPPLVTFEPLRCLDMK
HLEPDLPPVKTCTRLQFMLPKPSVMKRRQOCOREEDVIDLQGSWYKRLDDA
EUVVDRLLEHSGAEMTRQAVDAARGAIGDTGLDPVKSIGDTIGNYIVRVPPA
ETRYVHPSLEERYEPTPALDVEVETPTLPVYVMPSPSYEVEERDLRAVRYRMS
AAQAVLDCKHWKMGVGDSESDQLRFVEMQOPWEAEILIRMPDGEITVYVAGADM
MGBLIVSEHALRDTYFFEFBOAGLHGIAGEKMDPVMGLGARGDTIIVYAGADM
ETELRCHGGLDMVEKVCVCAQDDGERMIACDACDVHHTRCVGIADSEBVPPLFLC
VLCGALMAAGPINS"
3576..3612
/rpt_family="GC_rich"
3700..3732
/rpt_family="GC_rich"
3818..3857
/rpt_family="GC_rich"
6907..6929
/rpt_family="AT_rich"
complement(8447..11020)
/gene="OSJNBa0078A17.2"
/note="Predicted by fgenesh"
complement(join(<8447..8572,8747..8812,10002..10187,
10261..10798,10947..11020))
/gene="OSJNBa0078A17.2"
complement(join(8447..8572,8747..8812,10002..10187,
10261..10798,10947..11020))
/gene="OSJNBa0078A17.2"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA038820.1"
/db_xref="GI:28301927"
/translation="MADTTLRPSGDAAKKIADALRKTGVSNAVRTVIRVKGTYE
EILDKLEKILNLYKIVHSDDSTRTORAGHPRLPAHPSPVRKLVFVFNGRALKM
IERLRILKIASVATPSKMFILVSAVAVPAAVHNGNIVNIGRIGGGGGGGGGG
VGDDGGGGGPNVGG
WPRYLAVNKGSEKLVPRDGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
VKDAVPEGTSTISITLFAKKNVGAIVMDKEAPMAISPAISQYSTRSDPCF"
10295..10431
/rpt_family="CT_rich"
10295..10350
/rpt_family="rander"
12000..12251
/rpt_family="rander"
12005..12251
/rpt_family="rander"
12093..12120
/rpt_family="AT_rich"
12387..12419
/rpt_family="AT_rich"
13003..13066
/rpt_family="CT_rich"
13074..13156
/rpt_family="CGG n"
18206..18239
/rpt_family="rander"

gene complement(20076..20474)
/gene="OSJNBa0078A17.3"
/note="predicted by fgenesh"
complement(<20076..20474)
/gene="OSJNBa0078A17.3"
complement(20076..20474)
/gene="OSJNBa0078A17.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA038822.1"
/db_xref="GI:28301929"
/translation="MCPILSPSPSSPSLLATPAPSLPMPQAPAAAGAGDGERRR
GGGGRHWDGGLGNRIILAAASAFIYAVITAVLLVDTSNEMDELFSPPPTANLL
LRPFLVLMKRTHTGELNFSACRSCF"
20304..20397
/rpt_family="CCG n"
22377..22404
/rpt_family="AT_rich"
23620..23680
/rpt_family="AT_rich"
24283..24305
/rpt_family="GC_rich"
25474..25666
/rpt_family="CCG n"
25833..25973
/rpt_family="CCG n"
complement(28258..32088)
/gene="OSJNBa0078A17.4"
/note="nearly identical to gibberellin induced receptor
like kinase TKM GB:CA669028 GI:1707642 (Oryza sativa)
Plant Physiol 112, 1397-1397 (1996); EST D41588,
C22449, D22039, D22341, BE040330 from this gene"
complement(join(<28258..28982,29676..32088))
/gene="OSJNBa0078A17.4"
complement(join(28453..28982,29676..32088))
/gene="OSJNBa0078A17.4"
/codon_start=1
/product="gibberellin-induced receptor-like kinase TKM"
/protein_id="AA038825.1"
/db_xref="GI:28301932"
/translation="MMRRRRGGGGGGGGGVAVALVAVLWVYVAAAGETAASDAAM
RAYARALGADKALGMSGDPSCSPMAAGTQCSAGRVAVOVGNLSLGLRALPEVRN
LITALARLELDNLSIGELPSLAGLSIQLILVHNKFTPIPDFFGLTLRAVSLDN
NPDPWPLPADLADCTSLNFSANTAVTALPDFFGTALPSLQSLFANRMSGVP
ASLATALQALMLNNGENQFNGSISFTSNMSTLOELMHSIDFSLPDPFGSLASL
SDLEARNOLGVPDSILKLSLTVTLLNLLQPTPEKADKADAVPTTERFCL
STPGQPCDPRVNLLEVAEFQYPAKLADWKGNDODGYIGVGDAGNITVLNPFAM
GFGSISPAIGKITTLQKLIADNNITGTQKREVALPALTVDLSNNNLXGLPTPEA
AKNVLVANGSPNIGKDAFAPSGSGSGSMAFDGNGDGSNGSPSSSAGITAGSV
VGAIVAGLALAGFPYCKRKRPFGRVOSPAMVHPHRSQSDPVMKITYAGGVN
GAAASRTYQASGPRDILVAVETGMNVSIVLRVNTNFSDENLGGRGTYYVG
ELHGTGIAIKREMGAGVGNKGLPEWKRSLISALDVARGVYHSLAQOFTIHRDLKPS
EYMOGSLIATPEWKEHNRPLPEWKRSLISALDVARGVYHSLAQOFTIHRDLKPS
NIILGDMKAKVADPGLVRLAPADKCVSERTLAPGLAEVAVTGRVTTKADV

Query Match 14.1%; Score 70; DB 8; Length 147344;
Beet Local Similarity 100.0%; Pred. No. 7.4e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAAGAGAGTCGGGATAGTCCAAATTAACAAAGC 415
|||||
Db 86052 TACTCGAGTCATTCATATGCTTGAAGAGAGTCGGGATAGTCCAAATTAACAAAGC 86111
|||||
QY 416 TAAGATTACC 425
|||||
Db 86112 TAAGATTACC 86121
|||||

RESULT 14
150114 150114 1392 bp DNA linear PAT 07-OCT-1997
LOCUS 150114
DEFINITION Sequence 6 from patent US 5641876.
ACCESSION 150114

```

VERSION      150114.1  GI:2472334
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    Unclassified.
AUTHORS      1 (bases 1 to 1392)
TITLE        McElroy, D. and Wu, R.
JOURNAL      Rice actin gene and promoter
FEATURES     Patent: US 5641876-A 6 24-JUN-1997;
              Location/Qualifiers
                1..1392
                /organism="unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match      13.7%; Score 68; DB 6; Length 1392;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358  CTGAGGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 417
          |||||||
          1  CTGAGGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 60

DB      418  AGATTACC 425
          |||||||
          61  AGATTACC 68

RESULT 15
LOCUS       EVCOR112N      4412 bp      DNA      linear      SYN 18-MAR-1996
DEFINITION  Expression vector pCOR112N (modified from pCOR112 in [2]).
ACCESSION   Z29587
VERSION     Z29587.1  GI:452344
KEYWORDS    ampicillin resistance; beta-lactamase; colEI origin; expression
            vector; nos terminator; phage f1 region; rice actin1 promoter.
SOURCE      synthetic construct
            other sequences; artificial sequences.
REFERENCE   1 (bases 27 to 282)
AUTHORS     Depicker, A., Stachel, S., Dhaese, P., Zambrzyki, P. and Goodman, H.M.
TITLE       Nopaline synthase: transcript mapping and DNA sequence
JOURNAL     J. Mol. Appl. Genet. 1 (6), 561-573 (1982)
MEDLINE    83110651
PUBMED     7153689

REFERENCE   2
AUTHORS     McElroy, D., Blowers, A.D., Jones, B. and Wu, R.
TITLE       Construction of expression vectors based on the rice actin 1 (Act1)
JOURNAL     Mol. Gen. Genet. 231 (1), 150-160 (1991)
MEDLINE    92092956
PUBMED     1753941
REMARK      (sites)
REFERENCE   3 (bases 1 to 4412)
AUTHORS     Liu, L., Dasgupta, I., Davies, J.W. and Hull, R.
TITLE       Modified vectors for monocot transformation toward virus resistance
JOURNAL     Unpublished
REFERENCE   4 (bases 1 to 4412)
AUTHORS     Liu, L.
TITLE       Direct Submision
JOURNAL     Submitted (26-JAN-1994) Liu L., John Innes Institute, Virus
            Research, Colney Lane, Norwich, United Kingdom, NR4 7UH
            Location/Qualifiers
              1..4412
              /organism="synthetic construct"
              /mol_type="other DNA"
              /db_xref="taxon:32630"
              1..30
              /note="multiple cloning sites"
              3..5
              /note="start codon for gene expression"
              27..282
              /note="nos terminator"

```

```

misc_feature    /citation=[1]
                506..962
                /note="phage f1 region"
CDS             1093..1953
                /codon_start=1
                /transl_table=11
                /product="beta-lactamase"
                /protein_id="CAA82704.1"
                /db_xref="GI:452345"
                /translation="MSIQHRYVALIPFPAAFCPLPVPAPHPETLVKVKADQDGLGARVGY
                IELDINSKGLTSEFRPEERFPMWSTFKVLGCAYISRIDAGQEQGLRIRHYSQNDLVE
                VSPVTEKHLIDMTVRELCSAAITWSDNTAANLLITIGQPELTFLLNMGDVTPL
                DRWPELNAEIPNDRDITTPVMAATTLEKLLTGSLTLASQQLIDMEADKVAQPL
                LRSALPAGWFIADKSGAGRGSRGIIAALGPDKPSRIIVIYTTGSQATWDERNRQIA
                EIGASLIKHM"
                1954..2894
                /note="ColEI origin"
                3181..4412
                /note="rice actin 1 5' region (act1 promoter)"
                /citation=[2]

ORIGIN
Query Match      13.7%; Score 68; DB 12; Length 4412;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358  CTGAGGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 417
          |||||||
          DB      3181  CTGAGGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 3240

QY      418  AGATTACC 425
          |||||||
          DB      3241  AGATTACC 3248

Search completed: February 10, 2005, 13:13:39
Job time : 3075.76 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 111.027 Seconds
(without alignments)
7855.666 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18
Sequence: 1 tgcgtctcgtcgtacttc 18

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	6	AX342372 Sequence
2	18	100.0	1183	6	AX342369 Sequence
3	18	100.0	2386	10	BC055684 Mus muscu
4	18	100.0	2699	10	BC034572 Mus muscu
5	18	100.0	2835	10	BC027167 Mus muscu
6	18	100.0	3788	10	AY341877 Mus muscu
7	18	100.0	175440	10	AC125374 Mus muscu
8	18	100.0	218326	2	AC094355 Rattus no
9	17	94.4	140256	10	AC132091 Mus muscu
10	17	94.4	164405	9	AC016256 Homo sapi
11	17	94.4	172148	3	LMPP696
12	17	94.4	175062	10	AC138156 Mus muscu
13	17	94.4	229391	2	AC105823 Rattus no
14	17	94.4	231074	2	AC142509 Rattus no
15	17	94.4	231140	2	AC097799 Rattus no
16	17	94.4	233915	10	AC135290 Mus muscu
17	17	94.4	251289	2	AC095160 Rattus no
18	17	94.4	256017	2	AC121398 Rattus no
19	16	88.9	359	8	AY301027 Zea luxuri

20	16	88.9	359	8	AY301036 Zea diplo
21	16	88.9	365	8	AY301028 Zea may
22	16	88.9	365	8	AY301031 Zea may
23	16	88.9	365	8	AY301037 Zea diplo
24	16	88.9	371	8	AY301029 Zea may
25	16	88.9	371	8	AY301032 Zea may
26	16	88.9	371	8	AY301033 Zea may
27	16	88.9	372	8	AY301035 Zea peren
28	16	88.9	372	8	AY301030 Zea may
29	16	88.9	384	8	AY301034 Zea peren
30	16	88.9	559	8	AY296473 Zea may
31	16	88.9	559	8	AY296474 Zea may
32	16	88.9	559	8	AY296476 Zea may
33	16	88.9	559	8	AY296477 Zea may
34	16	88.9	559	8	AY296481 Zea may
35	16	88.9	559	8	AY296482 Zea may
36	16	88.9	560	8	AY296458 Zea may
37	16	88.9	561	8	AY296470 Zea may
38	16	88.9	562	8	AY296475 Zea may
39	16	88.9	562	8	AY296479 Zea may
40	16	88.9	562	8	AY296480 Zea may
41	16	88.9	562	8	AY296483 Zea may
42	16	88.9	563	8	AY296456 Zea may
43	16	88.9	563	8	AY296460 Zea may
44	16	88.9	563	8	AY296462 Zea may
45	16	88.9	563	8	AY296463 Zea may

ALIGNMENTS

RESULT 1
AX342372 18 bp DNA linear PAT 12-JAN-2002
LOCUS AX342372
DEFINITION Sequence 11 from Patent EP167531.
ACCESSION AX342372
VERSION AX342372.1 GI:18151815
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Behr, C.F., Hironaka, C., Heck, G.R. and You, J.
TITLE Corn transformant pv-zmgf32 (Ink603) and compositions and methods
for detection thereof
JOURNAL Patent: EP 1167531-A 11 02-JAN-2002;
Monsanto Technology LLC (US)
FEATURES
source
1..18
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Zea maize genomic DNA and vector DNA"
ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.8; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;
Cx 1 TGCTGTTCTGCTGACTTT 18
Db 1 TGCTGTTCTGCTGACTTT 18
RESULT 2
AX342369 1183 bp DNA linear PAT 12-JAN-2002
LOCUS AX342369
DEFINITION Sequence 8 from Patent EP167531.
ACCESSION AX342369
VERSION AX342369.1 GI:18151812
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct

other sequences; artificial sequences.

1
Behr, C.F., Hironaka, C., Heck, G.R. and You, J.
Corn transformant pv-zmg132 (nk603) and compositions and methods
for detection thereof
Patent: EP 1167531-A 8 02-JAN-2002;
Monsanto Technology LLC (US)
Location/Qualifiers
1. 1183
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="1-164 Agrobacterium tumefaciens nos 3' terminator
165-381 construct vector DNA 382-686 Zea mays genomic DNA"
genes, rps1 and rps4 687-1183 Zea mays genomic DNA"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1183;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
|||||
678 TGCTGTTCTGCTGACTTT 695

RESULT 3
BC055684 2386 bp mRNA linear ROD 12-OCT-2004
LOCUS Mus musculus zinc finger protein 445, mRNA (cDNA clone
IMAGE:6410700), complete cds.
BC055684
BC055684.1 GI:33416952

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
MUS musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2386)

REFERENCE
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Shabtay, T.E., Brownstein, M.J., Uedin, T.B., Topolyski, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Mulhally, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W.,
Vallat, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Sanchez, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2386)
Director MGC Project.
Direct Submission
Submitted (01-AUG-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) md@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 125 Row: 5 Column: 24
This clone has the following problem: The cds is short compared to
the longest cds in the locus.
Location/Qualifiers
1. 2386
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6410700"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="Vector: pTX-ASC"
1. 2386
/gene="Zfp445"
/note="synonym: ZNF168"
/db_xref="LOCUSID:235682"
/db_xref="WGI:2143340"
319. 1482
/gene="Zfp445"
/codon_start=1
/product="Zfp445 protein"
/protein_id="AAH55684.1"
/db_xref="GI:33416953"
/db_xref="LOCUSID:235682"
/db_xref="WGI:2143340"
/translation="MTKLHNEKVKQKROEDPFKQSYRQSGVISTVEKPECONCGKTF
TOKSLIEHORITGEKPYQSCGFTYRSYIIMRTQAIKKEHGLTFSSQ
GAVEPIPRGSHTEGSKYCGKAFHNSFLIHRVHTREKPYQRCREKAFRNS
NIVRHOQKFLHRYKREKSTNSQSLIDKQPMOCGKTFTRRSLLDHGCI
HSGRRKCNLCSEKSPRNLVNHORITTEPQMRDQPVGTHARSVDQRKSHVTL
QSEGLHSDPGSYSCDVNLINQESGKURKCCNPSSSSKIAFQVPTKAKCH
KCSYCGKTFKSHLSHKKCHTKRPFKICVCGKTFRWSNLTIRMKQHVNR"

ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 2386;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
|||||
1868 TGCTGTTCTGCTGACTTT 1885

RESULT 4
BC034572 2699 bp mRNA linear ROD 12-OCT-2004
LOCUS Mus musculus zinc finger protein 445, mRNA (cDNA clone
IMAGE:4018895), partial cds.
BC034572
BC034572.1 GI:21961272

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2699)

REFERENCE
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

[illegible][illegible]

JOURNAL Submitted (24-APR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 175440)
AUTHORS Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 24, 2003 this sequence version replaced gr:29501931.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_BA0438P19

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Oseegawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.reggen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES
Source location/Qualifiers

1..175440
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/map="g"
/clone="RP23-438P19"
/clone_id="RPCI-23"
repeat_region
/rpt_family="Alu"
637..781
/rpt_family="Alu"
871..1011
/rpt_family="B2"
repeat_region
/rpt_family="B2"
1421..1533
/rpt_family="L2"
repeat_region
/rpt_family="B4"
1813..1948
/rpt_family="B4"
1949..2002
/rpt_family="L2"
repeat_region
/rpt_family="L2"
2123..2286
/rpt_family="Alu"
2305..2481
/rpt_family="MER2_type"
repeat_region
/rpt_family="L1"
3024..3138
/rpt_family="L1"
3143..3318
/rpt_family="L1"
repeat_region

repeat_region 3603..3736
/rpt_family="Alu"
repeat_region 4222..4324
/rpt_family="L1"
repeat_region 4343..4428
/rpt_family="B4"
repeat_region 4742..4914
/rpt_family="B2"
repeat_region 5197..5282
/rpt_family="MER2_type"
repeat_region 5423..5493
/rpt_family="MER2_type"
repeat_region 7280..7429
/rpt_family="Alu"
repeat_region 7815..7964
/rpt_family="Alu"
repeat_region 8235..8348
/rpt_family="MIR"
repeat_region 8485..8621
/rpt_family="Alu"
repeat_region 8706..8770
/rpt_family="ID"
repeat_region 9283..9524
/rpt_family="L1"
repeat_region 9575..9914
/rpt_family="MaLR"
repeat_region 9992..10122
/rpt_family="Alu"
repeat_region 10127..10223
/rpt_family="U6"
repeat_region 10453..10584
/rpt_family="Alu"
repeat_region 10602..10689
/rpt_family="Alu"
repeat_region 11339..11456
/rpt_family="Alu"
repeat_region 12314..12451
/rpt_family="Alu"
repeat_region 12620..12770
/rpt_family="Alu"
repeat_region 13406..13586
/rpt_family="B2"
repeat_region 13598..13647
/rpt_family="L2"
repeat_region 13697..13891
/rpt_family="ERVX"
repeat_region 14688..14992
/rpt_family="RMR12"
repeat_region 15621..16001
/rpt_family="MaLR"
repeat_region 16163..16348
/rpt_family="RMR12"
repeat_region 16944..17521
/rpt_family="ERV1"
repeat_region 18760..18903
/rpt_family="Alu"
repeat_region 19583..19723
/rpt_family="Alu"
repeat_region 19764..19870
/rpt_family="B4"
repeat_region 20295..20771
/rpt_family="RMR6B"
repeat_region 20804..20976
/rpt_family="ERVX"
repeat_region 20974..21254
/rpt_family="ERVX"
repeat_region 21255..21323
/rpt_family="ERVX"
repeat_region 21324..21840
/rpt_family="ERVX"
repeat_region 21803..21884
/rpt_family="ERV1"
repeat_region 21901..22284

```

repeat_region      /rpt_family="ERVVK"
                    22726. .22884
                    /rpt_family="B4"
                    22860. .22923
                    /note="Sequence derived from one plasmid subclone."
unsure              22958. .23159
                    /rpt_family="B2"
                    23998. .24132
                    /rpt_family="Alu"
repeat_region      24532. .24592
                    /rpt_family="B2"
repeat_region      24937. .25150
                    /rpt_family="B4"
repeat_region      25176. .25353
                    /rpt_family="B2"
repeat_region      25356. .25438
                    /rpt_family="B4"
repeat_region      25914. .26000
                    /rpt_family="S5"
repeat_region      28142. .28283
                    /rpt_family="Alu"
repeat_region      28298. .28352
                    /rpt_family="B2"

```

```

Query Match      100.0%; Score 18; DB 10; Length 175440;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCCTTCTCTGCTGACTTT 18
Db 105579 TGCCTTCTCTGCTGACTTT 105596

```

```

RESULT 8
AC094355
LOCUS      AC094355      218326 bp      DNA      linear      HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-3C12, *** SEQUENCING IN PROGRESS ***
ACCESSION  AC094355
VERSION     AC094355.6 GI:30467252
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

```

REFERENCE
AUTHORS    1 (bases 1 to 218326)
            Muzny,D,Marle, Metzker,M, Lee, Abramson,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alibrooks,S., Amin,A., Angiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Cesaar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Dextamo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Fahley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorenzowa,L., Louisedge,H., Lozada,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmood,M., Malloy,K., Mangum,A.,

```

TITLE

JOORNAL

REFERENCE

JOORNAL

AUTHORS

JOORNAL

COMMENT

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milošević,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemelehen,O., Okwunonu,G., Olarinmusaogun,A., Pal,S., Parks,K.,
Paoletnak,S., Paul,H., Perez,A., Perez,L., Plankhock,C.,
Plopper,F., Polidexter,A., Popovic,D., Primus,E., Pu,L.,-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Rellly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Soosa,J.,
Steinle,M., Strong,R., Sutton,A., Swack,A., Tabors,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanji,K.,
Valas,R., Vera,V., Villaesna,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willison,R., Wlezyk,R., Wooden,H., Wortley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE

JOORNAL

JOORNAL

JOORNAL

REFERENCE

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

JOORNAL

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 218326)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GALT
Center clone name: CH230-3C12
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 199312 bases at least Q40
Consensus quality: 201330 bases at least Q30
Consensus quality: 202750 bases at least Q20
Estimated insert size: 207083; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 19397: contig of 19397 bp in length
* 19398 19497: gap of unknown length
* 19398 204378: contig of 184881 bp in length
* 204379 204478: gap of unknown length
* 204479 216833: contig of 12355 bp in length
* 216834 216933: gap of unknown length
* 216934 218326: contig of 1393 bp in length.

FEATURES
source Location/Qualifiers
1..218326
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3C12"
204479..206130
/note="wgs_contig"

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 218326;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18
|||||
Db 65609 TGCTGTTCTGCTGACTTT 65626

RESULT 9
AC132091/c 140256 bp DNA linear ROD 30-JUL-2004
LOCUS Mus musculus chromosome 18 clone RP24-46SH15, complete sequence.
DEFINITION AC132091
AC132091.3 GI:50839052
VERSION
KEYWORDS
HTG.
SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE
AUTHORS Wilson,R.K.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 140256)
Wilson,R.K.
REFERENCE
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 140256)
Wilson,R.K.
REFERENCE
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Jul 30, 2004 this sequence version replaced gi:34495013.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BB0465H15

Location/Qualifiers
1..140256
/organism="Mus musculus"
/mol_type="genomic DNA"

FEATURES
source

/db_xref="taxon:10090"
/chromosome="18"
/clone="RP24-46SH15"

ORIGIN

Query Match 94.4%; Score 17; DB 10; Length 140256;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCTGTTCTGCTGACTTT 18
|||||
Db 69580 GCTGTTCTGCTGACTTT 69564

RESULT 10
AC016256/c 164405 bp DNA linear PRI 25-JUL-2001
LOCUS Homo sapiens 12 BAC RP11-98B6 (Roswell Park Cancer Institute Human
BAC library) complete sequence.
DEFINITION AC016256
AC016256.19 GI:14670071
VERSION
KEYWORDS
HTG.
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 164405)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Aze,J.R., Banks,T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carion,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chaver,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emertling,S., Escotto,M., Falls,T., Ferraguto,D.,
Elmag,S., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gortelli,J.H., Guevara,M.,
Gunsarane,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homel,F.,
Howard,S., Huber,J., Hulik,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Koryah,J., Kovar,C.,
Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichteberg,O., Lieu,C., Liu,J., Liu,W.,
Loulleged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Marondei,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Maunthey,B., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokoko,S.,
Ogun,M., Okwoum,G., Otaguwa,N., Oriado,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pichens,R., Pitman,E., Pol,L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuboakan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shoohrtari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Swatek,A., Taber,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wiczak,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Zuchelapatti,R. and Gibbs,R.

JOURNAL Submitted (24-NOV-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 164405)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 164405)

Worley,K.C.

REFERENCE Direct Submission

AUTHORS Submitted (25-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL On Jul 11, 2001 this sequence version replaced gi:12656764.

COMMENT INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES

source

1.164405

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="12"

/clone="RP11-98B6"

/complement(1..371)

/rpt_family="MLT1F1"

3..5166

/note="Overlaps bases 186760..191923 of clone AC087886"

/function="Overlaps with adjacent clone AC087886"

repeat_region

/rpt_family="L1MA8"

2142..2281

/rpt_family="MIR"

repeat_region

/complement(2827..2961)

/rpt_family="MIR"

3943..4314

repeat_region

/rpt_family="L2"

repeat_region

5502..5527

/rpt_family="AT_rich"

repeat_region

6069..6615

/rpt_family="L1ME3A"

repeat_region

6830..6911

/rpt_family="L2"

repeat_region

7015..7068

/rpt_family="AT_rich"

repeat_region

7069..7153

/rpt_family="(TA)n"

repeat_region

7163..7191

/rpt_family="(TG)n"

repeat_region

complement(7192..7324)

/rpt_family="L2"

repeat_region

complement(7387..7463)

/rpt_family="L2"

repeat_region

7825..8933

/rpt_family="L1PA13"

repeat_region

complement(8937..9171)

/rpt_family="L2"

repeat_region

9494..9629

/rpt_family="L1"

repeat_region

complement(9700..10075)

/rpt_family="L1MA9"

repeat_region

10071..10234

/rpt_family="L1MA9"

repeat_region

10294..10581

/rpt_family="L1MA4"

repeat_region

10582..11009

/rpt_family="Tigger2a"

repeat_region

11010..11303

/rpt_family="L1M4"

repeat_region

11500..11670

/rpt_family="L1M3"

repeat_region

11701..12100

/rpt_family="L1M3"

repeat_region

12185..13027

/rpt_family="L1M3"

repeat_region

complement(13224..13505)

/rpt_family="L1PA10"

repeat_region

13829..13957

/rpt_family="AT_rich"

repeat_region

complement(13958..15549)

/rpt_family="L1PA3"

misc_feature

14218

/note="A or G"

/function="Polymorphic site"

14264

/note="C or A"

/function="Polymorphic site"

14283

/note="C or A"

/function="Polymorphic site"

14288

/note="C or T"

/function="Polymorphic site"

14300..14431

/standard_name="G19948"

/db_xref="dbSTS:32826"

14356

/note="C or T"

/function="Polymorphic site"

14362

/note="C or G"

/function="Polymorphic site"

14397..14398

/note="CA or TG"

/function="Polymorphic site"

14426

Query Match 94.4%; Score 17; DB 9; Length 164405;

Best local similarity 100.0%; Pred. No. 13;

	Matches	17; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	TGCTGTTCGTGAATT	17	:					
Db	51082	TGCTGTTCGTGAATT	51066						

clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is overlapped by AC131796 and AC117191.

FEATURES

source

1.175062
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/map="18"
/clone="RP24-115J20"
/clone_1b="RPCI-24"
2816.2832
/rpt_family="Alu"
2833.3199
/rpt_family="ERVK"
3200.3270
/rpt_family="Alu"
3284.4049
/rpt_family="L1"
4050.4441
/rpt_family="ERV1"
4589.4801
/rpt_family="L1"
4804.5251
/rpt_family="L1"
5267.6350
/rpt_family="L1"
6369.6443
/rpt_family="ERVK"
6607.6651
/rpt_family="ERVK"
7020.8857
/rpt_family="ERVK"
9542.9763
/rpt_family="ERVK"
9900.10153
/rpt_family="ERVK"
10154.10574
/rpt_family="ERVK"
10577.11619
/rpt_family="L1"
13136.13296
/rpt_family="L1"
13297.13771
/rpt_family="ERV1"
13772.14140
/rpt_family="L1"
15017.15329

/rpt_family="L1"
18263.18690
/rpt_family="ERVK"
19452.19708
/rpt_family="ERVK"
20128.20850
/rpt_family="ERVK"
21320.21462
/rpt_family="ERV1"
22527.22679
/rpt_family="ERV1"
22867.23274
/rpt_family="ERVK"
23506.23612
/rpt_family="ERVK"
23814.24356
/rpt_family="L1"
24451.25827
/rpt_family="L1"
25898.26118
/rpt_family="MALR"
28719.29029
/rpt_family="L1"
29051.30039
/rpt_family="L1"
30145.30382
/rpt_family="L1"
31201.31495
/rpt_family="B4"
31879.32154
/rpt_family="ERV1"
32431.32622
/rpt_family="B2"
33168.33241
/rpt_family="B2"
33781.33952
/rpt_family="B2"
34262.34352
/rpt_family="Alu"
34418.34886
/rpt_family="L1"
41760.42428
/rpt_family="L1"
42404.42444
/rpt_family="L1"
44749.44822
/rpt_family="MALR"
45324.45314
/rpt_family="B4"
45830.46019
/rpt_family="B2"
46075.46440
/rpt_family="MALR"
46551.46750
/rpt_family="B2"
46787.46896
/rpt_family="B4"
47633.47812
/rpt_family="B2"
47972.48062
/rpt_family="L1"
48820.49034
/rpt_family="B4"
49400.49484
/rpt_family="L1"
49700.49854
/rpt_family="Alu"
49948.50004
/rpt_family="L1"
50087.50185
/rpt_family="Alu"
50130.50186
/rpt_family="B4"


```

repeat_region      50322..50451
                    /rpt_family="MalaR"
repeat_region      51460..51517
                    /rpt_family="L1"
repeat_region      51610..51885

Query Match      94.4%; Score 17; DB 10; Length 175062;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GCGTCTCGCTGACTTT 18
Db      59764 GCGTCTCGCTGACTTT 59780

RESULT 13
AC105823/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-141P22, WORKING DRAFT SEQUENCE, 7
unordered pieces.
AC105823      229391 bp      DNA      linear      HTG 15-NOV-2002
AC105823
AC105823.4 GI:25007047
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229391)
Wozny,D,Marie, Metzker,M,Lee, Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bernahmed,F.,
Bismalo,K., Blair,J., Blankenburg,K., Blych,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacero,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devilla,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falle,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,J., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Gante,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunnarsson,P., Healand,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlerbird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,Z., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenzhuwa,L., Louisedge,H., Lozada,R.O., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwokoelameh,O., Okwuonu,G., Olairnusaagun,A., Pal,S., Parks,K.,
Paternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Rose,R., Ruiz,S.J., Riggs,F.,
Rives,C., Rodkey,T., Rojase,A., Rose,M., Rose,R., Ruiz,S.J., Riggs,F.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shattman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sned,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabors,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K.,
Valdes,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNVH
Center clone name: CH230-141P22
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 209637 bases at least Q40
Consensus quality: 212405 bases at least Q30
Consensus quality: 213820 bases at least Q20
Estimated insert size: 215551; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 10067: contig of 10067 bp in length
* 10068 10167: gap of unknown length
* 10168 30107: contig of 19940 bp in length
* 30108 30208: gap of unknown length
* 30208 219981: contig of 189774 bp in length
* 219981 220081: gap of unknown length
* 220081 224729: contig of 4648 bp in length
* 224729 224829: gap of unknown length
* 224829 225855: contig of 1026 bp in length
* 225855 225955: gap of unknown length
* 225955 227516: contig of 1561 bp in length
* 227516 227517: gap of unknown length
* 227517 227616: gap of unknown length
* 227616 229391: contig of 1775 bp in length.
*
* Location/Qualifiers
* 1..229391

```

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-141P22"

misc_feature
    1..1175
    /note="wgs_contig"

misc_feature
    8319..11067
    /note="wgs_contig"

misc_feature
    10168..11538
    /note="wgs_contig"

misc_feature
    22082..221295
    /note="wgs_contig"

ORIGIN

Query Match      94.4%; Score 17; DB 2; Length 229391;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGCTGTTCTGCTGACTT 17
Db      117028 TGCTGTTCTGCTGACTT 117012

RESULT 14
AC142509
LOCUS      AC142509
DEFINITION Rattus norvegicus clone CH230-unknown, WORKING DRAFT SEQUENCE, 38
            unordered pieces.
AC142509
VERSION    AC142509.1 GI:29501816
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 231074)
            Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
            Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D,
            Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
            Baldwin,D, Bandaruaike,D, Barber,M, Barnstead,M, Benahmed,F,
            Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
            Bryant,N, Buhay,C, Burch,P, Butrell,K, Calderon,E,
            Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Celler,A,
            Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
            Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
            Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
            Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
            Draper,H, Dugan-Rocha,S, Dunn,A, Dublin,K, Duval,B, Eaves,K,
            Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
            Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
            Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
            Gebregorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
            Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
            Harvery,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
            Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
            Hollins,B, Howells,S, Huliy,S, Hume,J, Idlebird,D, Jackson,A,
            Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
            Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,L,
            Kowis,C, Kraft,C,L, Lebow,H, Levan,U, Lewis,L, Li,Z, Liu,J,
            Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
            Lorenshuwa,L, Louisedge,H, Lozado,R,J, Lu,X, Ma,J,
            Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,
            Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
            Mawhinney,S, McLeod,M, McNeill,T, Meenen,E, Milosavljevic,A,
            Miner,G, Minja,E, Montemayor,U, Moore,S, Morgan,M, Morris,K,
            Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D,
            Newton,N, Nguyen,N, Norris,S, Nwaokemele,O, Okwona,G,
            Olariunsaogoon,A, Pal,S, Parks,K, Pasternak,S, Paul,H,
            Perez,A, Perez,L, Pfankoch,C, Plopper,F, Poldexter,A,
            Popovic,D, Primus,E, Pu,L-L, Puazo,M, Quiroz,J, Rachlin,E,
            Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y,
            Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A,

```

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

```

Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S.,
Scott,G., Shatman,S., Shen,H., Shetty,J., Shwartsbeyn,A.,
Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A.,
Swatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Umani,K., Valae,R., Vera,V., Villaana,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
Warren,R., Wei,X., White,F., Williams,G., Wilson,R., Wleczky,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yan,J., Yeon,L., Yeon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 231074)
Moriy,K.C.
Submitted (03-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KZZY
Center clone name: CH230-unknown
----- Summary Statistics
Sequencing vector: Plasmid,
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 216536 bases at least Q40
Consensus quality: 221871 bases at least Q30
Consensus quality: 225261 bases at least Q20
Estimated insert size: 214970, sum-of-coverage estimation
Quality coverage: 5x in Q20 bases, sum-of-coverage estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 38 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1537: contig of 1537 bp in length
1538 1637: gap of unknown length
1638 2663: contig of 1026 bp in length
2664 2764: gap of unknown length
2764 3850: contig of 1087 bp in length
3851 3950: gap of unknown length
3951 5198: contig of 1248 bp in length
5199 5298: gap of unknown length
5299 6650: contig of 1352 bp in length
6651 6750: gap of unknown length
6751 7852: contig of 1102 bp in length
7853 7952: gap of unknown length
7953 9555: contig of 1603 bp in length
9556 9656: gap of unknown length
9656 11097: contig of 1442 bp in length
11098 11197: gap of unknown length
11198 12750: contig of 1553 bp in length
12751 12850: gap of unknown length
12851 14233: contig of 1383 bp in length
14234 14333: gap of unknown length
14334 15578: contig of 1245 bp in length
15579 15679: gap of unknown length
15679 16807: contig of 1129 bp in length
16808 16907: gap of unknown length
16908 18834: contig of 1927 bp in length

```

18935 18934: gap of unknown length
 18935 20671: contig of 1731 bp in length
 20672 20771: gap of unknown length
 20772 24401: contig of 3630 bp in length
 24402 24501: gap of unknown length
 24502 26538: contig of 2037 bp in length
 26539 26638: gap of unknown length
 26639 30516: contig of 3878 bp in length
 30517 30616: gap of unknown length
 30617 34896: contig of 4280 bp in length
 34897 39921: contig of 4925 bp in length
 39922 40022: gap of unknown length
 40022 46416: contig of 6395 bp in length
 46417 46516: gap of unknown length
 46517 52171: contig of 5655 bp in length
 52172 52271: gap of unknown length
 52272 57654: contig of 5383 bp in length
 57655 57754: gap of unknown length
 57755 65163: contig of 7409 bp in length
 65164 65263: gap of unknown length
 65264 71371: contig of 6108 bp in length
 71372 71471: gap of unknown length
 71472 77084: contig of 5613 bp in length
 77085 77184: gap of unknown length
 77185 84572: contig of 7388 bp in length
 84573 84672: gap of unknown length
 84673 91928: contig of 7256 bp in length
 91929 92028: gap of unknown length
 92029 97903: contig of 5875 bp in length
 97904 98003: gap of unknown length
 98004 107470: contig of 9467 bp in length
 107471 107570: gap of unknown length
 107571 117528: contig of 9958 bp in length
 117529 117628: gap of unknown length
 117629 126174: contig of 8546 bp in length
 126175 126274: gap of unknown length
 126275 135616: contig of 9342 bp in length
 135617 144781: contig of 9065 bp in length
 144782 144881: gap of unknown length
 144882 158599: contig of 13718 bp in length
 158600 158699: gap of unknown length
 158700 175941: contig of 17242 bp in length
 175942 176041: gap of unknown length
 176042 194389: contig of 18348 bp in length
 194390 194489: gap of unknown length
 194490 213045: contig of 18556 bp in length
 213046 213145: gap of unknown length
 213146 231074: contig of 17929 bp in length.
 Location/Qualifiers
 1. 231074
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-unknown"

ORIGIN

Query Match 94.4% Score 17: DB 2: Length 231074;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTT 17
 |||||
 Db 206129 TGCTGTTCTGCTGACTT 206145

FEATURES

RESULT 15
 AC097799/c 231140 bp DNA linear HTG 10-MAY-2003
 LOCUS Rattus norvegicus clone CH230-46A18, WORKING DRAFT SEQUENCE, 2
 DEFINITION unordered pieces.
 ACCESSION AC097799

VERSION
 AC097799.7 GI:30521368
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 Rattus norvegicus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 231140)
 Muzny,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J,
 Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
 Anyalebech,V, Ayagi,A, Ayodeji,M, Baca,E, Baden,H,
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
 Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
 Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
 Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A,
 Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
 Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
 Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedertich,D,
 Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
 Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Davies,K,
 Egan,A, Escoto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
 Fernandez,S, Finley,M, Flegg,N, Forbes,L, Foster,M, Foster,P,
 Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
 Gebregeorgie,E, Geer,K, Gill,R, Grady,M, Guerra,M, Guevara,W,
 Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
 Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
 Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
 Hollins,B, Howells,S, Huylk,S, Hume,J, Idlebird,D, Jackson,A,
 Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
 Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
 Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
 Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
 Lorensewala,L, Louised,H, Lozada,R,J, Lu,X, Ma,J,
 Maheshwari,M, Mahindartne,M, Mahmoud,M, Malloy,K, Mangum,A,
 Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E,
 Mahoney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
 Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
 Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L,
 Nankervils,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Parks,K,
 Naoketleleh,O, Okunolu,G, Olarinuagbon,A, Pal,S, Parks,K,
 Pasternak,S, Paul,H, Perez,A, Perez,L, Plannkoeh,C,
 Plopper,F, Polndexter,A, Popovic,D, Primus,E, Pu,L,
 Piazto,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
 Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
 Rives,C, Rodkey,T, Rojase,A, Rose,M, Rose,R, Ruiz,S,J,
 Sanders,M, Saverly,G, Scherer,S, Scott,G, Shasman,S, Shen,H,
 Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smaja,D,
 Sneed,A, Sodergren,E, Song,X,Z, Sorelle,R, Sosa,J,
 Steidle,M, Strong,R, Sutton,A, Swalek,A, Taber,P, Taylor,C,
 Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K,
 Valas,R, Vera,V, Villaseña,D, Waldron,L, Walker,B, Wang,J,
 Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
 Williams,G, Willson,R, Wleczyk,R, Woodson,H, Worley,K,
 Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,
 Yu,F, Zhang,J, Zhou,J, Zhou,S, Zhao,S, Zhou,D, von
 Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
 Weinstein,G, and Gibbs,R.A.

TITLE
 JOURNAL
 REFERENCE
 2 (bases 1 to 231140)
 Worley,K.C.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231140)
 REFERENCE
 Rat Genome Sequencing Consortium.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced gi:24819446.
 The sequence in this assembly is a combination of BAC based reads

COMMENT
 The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GFGG
Center clone name: CH230-46A18

----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 218455 bases at least Q40
Consensus quality: 220324 bases at least Q30
Consensus quality: 222143 bases at least Q20
Estimated insert size: 231276; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length.
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 228436: contig of 228436 bp in length
* 228437 228536: gap of unknown length
* 228537 231140: contig of 2604 bp in length.

FEATURES

source

1. 231140

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-46A18"

misc_feature

1. 1042

/note="wgs_end_extension"

misc_feature

clone_end:T7"

5228. 8197

/note="wgs_end_extension"

misc_feature

clone_end:T7"

/note="clone_boundary"

clone_end:T7"

site:ECORI

misc_feature

end_sequence:BH328184"

complement(227335..228129)

/note="clone_boundary"

clone_end:Sp6

site:ECORI

end_sequence:BH328185"

ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 231140;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTT 17

DB 34562 TGCTGTTCTGCTGACTT 34546

Search completed: February 10, 2005, 13:13:54
Job time : 114.027 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 3366.47 Seconds
(without alignments)
5630.828 Million cell updates/sec

Title: US-10-790-430-7

Perfect score: 498
Sequence: 1 aatcgcattccaatcgcgac.....aaaagtcgccaagttaa 498

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	13.5	670	6	CB867772 HC01D06w
2	50	10.0	598	6	CB862324 HH05G12Y
3	34	6.8	793	9	CL786064 OR_BBA010
4	29	5.8	539	8	CC046788 3591.1.19
5	26	5.2	733	9	CL858930 OR_CBA009
6	25	5.0	893	9	CL795021 OR_CBA000
7	23	4.6	576	8	AO689332 nbdx0079G
8	23	4.6	654	8	AO255572 nbdx0011D
9	23	4.6	745	8	AO326905 OR_BBA000
10	23	4.6	778	8	AO326905 nbdx0080M
11	23	4.6	778	8	AO689711 nbdx0080M
12	23	4.6	850	8	AO573760 nbdx0080F
13	23	4.6	854	8	AO504264 OP_Ba001
14	22	4.4	601	9	CW523527 OP_Ba004
15	22	4.4	702	9	CL812779 OR_CBA002
16	22	4.4	713	9	CG002331 ZUDU030TV
17	22	4.4	733	9	CL858930 OR_CBA009
18	22	4.4	784	9	CL720122 OR_BBA004
19	22	4.4	801	9	CL811054 OR_CBA002
20	22	4.4	893	9	CL795021 OR_CBA000
21	21	4.2	654	9	CL615769 OR_BBA000
22	21	4.2	680	9	CL194957 104.419.1
23	21	4.2	692	8	AZ600338 1M0418C05
24	21	4.2	702	9	CL812779 OR_CBA002

25	21	4.2	738	9	CL832021 OR_CBA005
26	21	4.2	746	8	AO870229 nbe0036N
27	21	4.2	786	9	AG475192 Mus muscu
28	21	4.2	801	9	CL811054 OR_CBA002
29	21	4.2	811	8	CC018299 PUDK85TD
30	21	4.2	854	8	CC004665 PUDK11TD
31	21	4.2	1348	2	BE967273 601661142
32	20	4.0	276	8	B245511 CH230-314
33	20	4.0	352	8	CC383351 PUBIO15TD
34	20	4.0	386	9	CC702094 OCUF058TH
35	20	4.0	452	9	AG206444 Oryza sat
36	20	4.0	587	9	AO921459 RPI1-23-2
37	20	4.0	620	8	BH11311 RPI1-24-3
38	20	4.0	645	8	CG065648 PUJBR28TB
39	20	4.0	657	8	B2126584 CH230-490
40	20	4.0	657	9	CG913725 ZMMBBD037
41	20	4.0	661	9	CG065651 PUJBR28TD
42	20	4.0	686	2	BF274572 GA_ED002
43	20	4.0	699	9	CR155347 Reverse 8
44	20	4.0	720	9	AG575889 Mus muscu
45	20	4.0	730	5	BM932647 UI-M-CG0P

ALIGNMENTS

RESULT 1
CB867772
LOCUS
DEFINITION
CB867772 CH Hordeum vulgare cDNA clone HC01D06 5-PRIME, mRNA
sequence.

ACCESSION
CB867772
VERSION
CB867772.1 GI:30069748

KEYWORDS
SOURCE
ORGANISM
EST.
Hordeum vulgare
Hordeum vulgare

REFERENCE
AUTHORS
Granger, A.
Varshney, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and
1 (bases 1 to 670)
Pooideae; Triticeae; Hordeum.

TITLE
JOURNAL
COMMENT
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 670 Std Error: 0.00
Plate: 1 row: D column: 6
Seq primer: T7.

FEATURES

source
Location/Qualifiers
1..670
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cuiivar="Sloop"
/db_xref="GABI:511307"
/db_xref="taxon:4513"
/clone="HC01D06"
/tissue_type="coleoptile"
/dev_stage="coleoptile, 3 days old"
/lab_host="DH10B"
/clone_lib="CH"
/note="Vector: pSPOT; Site 1: SalI (5-end of cDNA);
Site 2: NotI (3-end of cDNA). Due to the cloning system
used blue/white selection for recombinants is not 100%
reliable. Average insert size is 1.3 kb."

ORIGIN

Query Match 13.5%, Score 67; DB 6; Length 670;

Best Local Similarity 100.0%; Pred. No. 4.3e-26;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAAACAAGCTAA 418
|||||
DB 58 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAAACAAGCTAA 117
|||||

QY 419 GATTACC 425
|||||
DB 118 GATTACC 124

RESULT 2
CB862324 598 bp mRNA linear EST 22-APR-2003
LOCUS CB862324
DEFINITION HH05G12Y HH Hordeum vulgare cDNA clone HH05G12 3-PRIME, mRNA
ACCESSION CB862324
VERSION CB862324
KEYWORDS
SOURCE EST.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 598)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graneir,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 598 Std Error: 0.00
Plate: 5 row: G column: 12
Seq primer: SP6.
Location/Qualifiers
1..598
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cullivar="Sloop"
/db_xref="GABI:533740"
/db_xref="taxon:4513"
/clone="HH05G12"
/issue_type="coleoptile"
/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site_1: SalI (5-end of cDNA);
Site_2: NciI (3-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 Kb."

ORIGIN
Query Match 10.0%; Score 50; DB 6; Length 598;
Best Local Similarity 100.0%; Pred. No. 2e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAA 408
|||||
DB 21 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAA 70
|||||

RESULT 3
CL786064 793 bp DNA linear GSS 30-JUL-2004
LOCUS CL786064
DEFINITION OR_BB80103K12.f OR_BBa Oryza rufipogon genomic clone OR_BB80103K12
5', genomic survey sequence.
ACCESSION CL786064

VERSION CL786064.1 GI:50865715
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 793)
Klim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
Omap Project
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA GG
Insert Length: 161 Std Error: 0.00
Plate: 0103 row: K column: 12
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers
1..793
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_BB80103K12"
/issue_type="young leaves"
/lab_host="DH10B-T1 phage resistant"
/clone_lib="OR_BBa"
/note="Vector: pGIRAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 6.8%; Score 34; DB 9; Length 793;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGT 389
|||||
DB 720 TACTCGAGTCATTCATATGCTTGAGAGAGAGT 753
|||||

RESULT 4
CC046788/c 539 bp DNA linear GSS 01-APR-2003
LOCUS CC046788/c
DEFINITION 3591.1 190.1 F04.Y.1 3591 - Rescuedu Gr-d P Zea mays genomic,
genomic survey sequence.
ACCESSION CC046788
VERSION CC046788.1 GI:29461679
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 539)
Walbot,V.
Maize genomic sequences found using engineered Rescuedu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2237
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 3591_1 190_1 row: 33
Class: transposon-tagged.

FEATURES
source
Location/Qualifiers
1. .539
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/873/K55"
/db_xref="taxon:4577"
/issue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="3591 - Rescuedu Grid P"
/note="Organ: leaf; Vector: Rescuedu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; Rescuedu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuedu, go to the web site 'www.zmldb.iastate.edu' and follow the links for 'Rescuedu', Grid P was grown at Molokai in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN
Query Match 5.8%; Score 29; DB 8; Length 539;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 CTGAGACGTGCGTCCCTGTGGAGCTGTCT 281
|||||
Db 443 CTGAGACGTGCGTCCCTGTGGAGCTGTCT 415

RESULT 5
CL858930/c 733 bp DNA linear GSS 16-AUG-2004
LOCUS OR_CBA0090A04.r OR_CBA Oryza rufipogon genomic clone OR_CBA0090A04
DEFINITION 3', genomic survey sequence.
ACCESSION CL858930
VERSION CL858930.1 GI:51270762
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Bharatoidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 733)
Kim,H., Yu,Y., Wiscocki,M., Yost,D., Stum,D., Rao,K., Luo,M., Jettly,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
J. J. Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0090 row: A column: 04
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.
Location/Qualifiers
1. .733
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA0090A04"
/issue_type="young leaves"
/dev_stage="2 week old seedlings"
/dev_stage="2 week old seedlings"

ORIGIN
Query Match 5.2%; Score 26; DB 9; Length 893;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 ATCGTAATAAAGTGCCCAAGT 495
|||||
Db 365 ATCGTAATAAAGTGCCCAAGT 340

RESULT 6
CL795021/c 893 bp DNA linear GSS 06-AUG-2004
LOCUS OR_CBA004L03.r OR_CBA Oryza rufipogon genomic clone OR_CBA004L03
DEFINITION 3', genomic survey sequence.
ACCESSION CL795021 GI:51017037
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Bharatoidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 893)
Kim,H., Yu,Y., Wiscocki,M., Yost,D., Stum,D., Rao,K., Luo,M., Jettly,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
J. J. Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0004 row: L column: 03
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.
Location/Qualifiers
1. .893
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA004L03"
/issue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_CBA"
/note="Vector: pGIBAC1; Site 1: HindIII; Site 2: HindIII; dirk treated 36 hrs before harvest"

ORIGIN
Query Match 5.2%; Score 26; DB 9; Length 893;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 ATCGTAATAAAGTGCCCAAGT 495
|||||
Db 365 ATCGTAATAAAGTGCCCAAGT 340

RESULT 7
AO68933/c 576 bp DNA linear GSS 01-JUL-1999
LOCUS AO689332

DEFINITION nbxb0079612f CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0079612f, genomic survey sequence.

ACCESSION AQ689332

VERSION AQ689332.1 GI:5330500

KEYWORDS GSS.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1 (bases 1 to 576)

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: TAATACGACTCACAATAGCG

Class: BAC ends

High quality sequence stop: 231.

Location/Qualifiers

1. 576

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="nbxb0079612f"

/rissue_type="Leaf"

/lab_host="E. coli DH10B"

/clone_lib="CUGI Rice BAC Library"

/note="Vector: pBelBAC11, Site_1: HindIII, Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 5.0%; Score 25; DB 8; Length 576;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 474 GTATATAAGGTGGCCCAAGTGA 498
|||||
Db 93 GTATATAAGGTGGCCCAAGTGA 69

RESULT 8

LOCUS AQ255572

DEFINITION nbxb0011D02r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0011D02r, genomic survey sequence.

ACCESSION AQ255572

VERSION AQ255572.1 GI:3780054

KEYWORDS GSS.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1 (bases 1 to 617)

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: GGAACACGCTATGACATG

Class: BAC ends

High quality sequence start: 11

High quality sequence stop: 296.

Location/Qualifiers

1. 617

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="nbxb0011D02r"

/rissue_type="Leaf"

/lab_host="E. coli DH10B"

/clone_lib="CUGI Rice BAC Library"

/note="Vector: pBelBAC11, Site_1: HindIII, Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 617;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 371 ATATGCTTGAGAGAGAGTGGG 393
|||||
Db 309 ATATGCTTGAGAGAGTGGG 331

RESULT 9

LOCUS CL615769

DEFINITION OR_BBao008D16.r OR_BB A Oryza rufipogon genomic clone OR_BBao008D16 3', genomic survey sequence.

ACCESSION CL615769

VERSION CL615769.1 GI:49104999

KEYWORDS GSS.

SOURCE Oryza rufipogon

ORGANISM Oryza rufipogon

REFERENCE 1 (bases 1 to 617)

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: GGAACACGCTATGACATG

Class: BAC ends

High quality sequence start: 11

High quality sequence stop: 296.

Location/Qualifiers

1. 617

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="nbxb0011D02r"

/rissue_type="Leaf"

/lab_host="E. coli DH10B"

/clone_lib="CUGI Rice BAC Library"

/note="Vector: pBelBAC11, Site_1: HindIII, Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Ehrhartoideae; Oryzae; Oryza.
1 (bases 1 to 654)
Kim, H., Yu, Y., Stum, D., Yost, D., Rao, K., Luo, M., Jettly, R.,
Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.
OMAP Project
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert length: 161 Std Error: 0.00
Plate: 0008 row: D column: 16
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES
source

location/Qualifiers
1..654
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_BA0008D16"
/issue_type="young leaves"
/lab_host="DH10B-T1 phage resistant"
/clone_lib="OR_BA"
/note="Vector: pNG1BAC1, Site_1: HindIII, Site_2: HindIII"

ORIGIN

Query Match 4.6%; Score 23; DB 9; Length 654;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATATGCTTGAGAGAGAGTCGGG 393
|||||
Db 311 ATATGCTTGAGAGAGAGTCGGG 333

RESULT 10
AQ326905 745 bp DNA linear GSS 08-JAN-1999
LOCUS nbxb0039P05r CUGI Rice BAC library Oryza sativa (japonica
DEFINITION cultivar-group) genomic clone nbxb0039P05r, genomic survey
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AQ326905.1 GI:4118755
GSS.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 745)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Seq primer: GGAAACAGCTATGACCAGT
Class: BAC ends
High quality sequence start: 4
High quality sequence stop: 316.

FEATURES
source

1..745
/organism="Oryza sativa (japonica cultivar-group)"

ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 745;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATATGCTTGAGAGAGAGTCGGG 393
|||||
Db 309 ATATGCTTGAGAGAGAGTCGGG 331

RESULT 11
AQ689711 778 bp DNA linear GSS 01-JUL-1999
LOCUS nbxb0080M19f CUGI Rice BAC library Oryza sativa (japonica
DEFINITION cultivar-group) genomic clone nbxb0080M19f, genomic survey
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AQ689711.1 GI:5330879
GSS.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 778)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Seq primer: TTAATACAGCTCATTATGCGG
Class: BAC ends
High quality sequence stop: 279.

FEATURES
source

location/Qualifiers
1..778
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0080M19f"
/issue_type="leaf"

/lab_host="E. coli DH10B"
/clone_lib="CGI Rice BAC Library"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from *Oryza sativa*, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9%. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 371 ATATGCTTGAGAGAGAGTCGGG 393

DB 309 ATATGCTTGAGAGAGAGTCGGG 331

RESULT 12

AO573760 850 bp DNA linear GSS 02-JUN-1999
LOCUS nbxb0084f24r CUCI Rice BAC Library *Oryza sativa* (japonica
cultivar-group) genomic clone nbxb0084f24r, genomic survey
sequence.

ACCESSION AO573760
VERSION AO573760
KEYWORDS GSS.
SOURCE AO573760.1 GI:4974545

ORGANISM

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; *Oryza*.

REFERENCE

1 (bases 1 to 850)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACAGCATATGACCATG
Class: BAC ends
High quality sequence stop: 310.
Location/Qualifiers

FEATURES

source

1. .850
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0084f24r"
/issue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CGI Rice BAC Library"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those

inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from *Oryza sativa*, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9%. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 850;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 371 ATATGCTTGAGAGAGAGTCGGG 393

DB 309 ATATGCTTGAGAGAGAGTCGGG 331

RESULT 13

CM504264 854 bp DNA linear GSS 06-OCT-2004
LOCUS OP_Ba0015P11.f OP_Ba *Oryza punctata* genomic clone OP_Ba0015P11
5', genomic survey sequence.

ACCESSION CM504264
VERSION CM504264
KEYWORDS GSS.
SOURCE CM504264.1 GI:53833769

ORGANISM

Oryza punctata
Oryza punctata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; *Oryza*.

REFERENCE

1 (bases 1 to 854)
SanMiguel, P., Westerman, R., Kim, H., Yu, Y., Wisotski, M., Yost, D.,
Stum, D., Rao, K., Luo, M., Jecely, R., Kudrna, D., Miller, C.,
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S.A.
OMAP Project - Purdue University
Unpublished (2004)
Contact: Scott A. Jackson
Jackson Laboratory
Purdue University
915 W. State St., West Lafayette, IN 47907, USA
Tel: 7654963621
Fax: 7654967255
Email: sjackson@purdue.edu

COMMENT

Basecalling by phred version 0.020425.c. This sequence was derived
from the raw sequence read by clipping with Lucy version 1.19s.
Bases 27-880 of the raw sequence (length 1189) were retained after
clipping.
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 161000 Std Error: 0.00
Plate: 0015 row: P column: 11
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers

FEATURES

source

1. .854
/organism="Oryza punctata"
/mol_type="genomic DNA"
/db_xref="taxon:4537"
/clone="OP_Ba0015P11"
/issue_type="young leaves"
/lab_host="DH10B-T1 phage resistant"
/clone_lib="OP_Ba"

ORIGIN /note="Vector: PAGIBAC1; Site_1: HindIII; Site_2: HindIII"

Query Match 4.6%; Score 23; DB 9; Length 854;
Best Local Similarity 100.0%; Pred. No. 0.38;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 460 AAGTAAATATCGGTATATAA 482
DB 350 AAGTAAATATCGGTATATAA 328

RESULT 14
CM523527/c 601 bp DNA linear GSS 08-OCT-2004
LOCUS OP_Ba0049120.r OP_Ba Oryza punctata genomic clone OP_Ba0049120
DEFINITION 3' genomic survey sequence.

ACCESSION CM523527 GI:53997749
VERSION CM523527.1
KEYWORDS GSS.
SOURCE Oryza punctata
ORGANISM Oryza punctata

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
1 (bases 1 to 601)

REFERENCE Sami, D., Rao, K., Luo, M., Jettly, R., Kudrna, D., Muller, C.,
Hartfield, J., Soderlund, C., Wing, R. and Jackson, S.A.
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S.A.

REFERENCE OMAP Project - Purdue University
Unpublished (2004)
JOURNAL Contact: Scott A. Jackson
COMMENT Jackson Laboratory

Purdue University
915 W. State St., West Lafayette, IN 47907, USA
Tel: 765/4963621
Fax: 765/4967255

Email: sjackson@purdue.edu
Basecalling by phred version 0.020425.c. This sequence was derived
from the raw sequence read by clipping with licy version 1.198.
Bases 55-655 of the raw sequence (length 1219) were retained after
clipping.

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert length: 161000 Std Error: 0.00
Plate: 0049 row: 1 column: 20
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES Location/Qualifiers

1..601
/organism="Oryza punctata"
/mol_type="genomic DNA"
/db_xref="taxon:4537"
/clone="OP_Ba0049120"
/issue_type="young leaves"
/lab_host="DH10B-T1 phage resistant"
/clone_1="OP_Ba"
/note="Vector: PAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 4.4%; Score 22; DB 9; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 TCATTCAATGCTTGAGAGAG 386
DB 261 TCATTCAATGCTTGAGAGAG 240

RESULT 15
CL812779 702 bp DNA linear GSS 09-AUG-2004
LOCUS

DEFINITION OR_CBA0028M15.r OR_CBA Oryza rufipogon genomic clone OR_CBA0028M15
3' genomic survey sequence.

ACCESSION CL812779 GI:51054117
VERSION CL812779.1
KEYWORDS GSS.

SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
1 (bases 1 to 702)

REFERENCE Kim, H., Yu, Y., Wiscocki, M., Yost, D., Stum, D., Rao, K., Luo, M.,
Jettly, R., Kudrna, D., Muller, C., Hartfield, J., Soderlund, C. and
Wing, R.

REFERENCE OMAP Project
Unpublished (2004)
JOURNAL Contact: Rod A. Wing
COMMENT Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0028 row: M column: 15
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES Location/Qualifiers

1..702
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA0028M15"
/issue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B-T1 phage resistant"
/clone_1="OR_CBA"
/note="Vector: PAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

ORIGIN

Query Match 4.4%; Score 22; DB 9; Length 702;
Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATATGCTTGAGAGAGATCGG 392
DB 310 ATATGCTTGAGAGAGATCGG 331

Search completed: February 10, 2005, 17:01:45
Job time : 3370.47 secs

This Page Blank (uspto)

QY	181	GAGTAAATTTCACTTGGGCCACCTTTTATTAACGATATTTTACTTTATACACCTTTTAA	240
QY <td>181</td> <td>GAGTAAATTTCACTTGGGCCACCTTTTATTAACGATATTTTACTTTATACACCTTTTAA</td> <td>240</td>	181	GAGTAAATTTCACTTGGGCCACCTTTTATTAACGATATTTTACTTTATACACCTTTTAA	240
QY <td>241</td> <td>ACTGATGTTTTCACTTTTGACCGAGTAATCTTACCTTGTGTTTATTTTGAAGTAATCCGA</td> <td>300</td>	241	ACTGATGTTTTCACTTTTGACCGAGTAATCTTACCTTGTGTTTATTTTGAAGTAATCCGA	300
QY <td>241</td> <td>ACTGATGTTTTCACTTTTGACCGAGTAATCTTACCTTGTGTTTATTTTGAAGTAATCCGA</td> <td>300</td>	241	ACTGATGTTTTCACTTTTGACCGAGTAATCTTACCTTGTGTTTATTTTGAAGTAATCCGA	300
QY <td>301</td> <td>CTCTCTTTCACAGCATATGATGACCTTGAAGTAACTTGTATTAACGGGCGCCCTTAAGGA</td> <td>360</td>	301	CTCTCTTTCACAGCATATGATGACCTTGAAGTAACTTGTATTAACGGGCGCCCTTAAGGA	360
QY <td>301</td> <td>CTCTCTTTCACAGCATATGATGACCTTGAAGTAACTTGTATTAACGGGCGCCCTTAAGGA</td> <td>360</td>	301	CTCTCTTTCACAGCATATGATGACCTTGAAGTAACTTGTATTAACGGGCGCCCTTAAGGA	360
QY <td>361</td> <td>TATCAAGCTGGTACACGGCGACACACTTCCACTAGTGTGAGTGATCCGTATATC</td> <td>420</td>	361	TATCAAGCTGGTACACGGCGACACACTTCCACTAGTGTGAGTGATCCGTATATC	420
QY <td>361</td> <td>TATCAAGCTGGTACACGGCGACACACTTCCACTAGTGTGAGTGATCCGTATATC</td> <td>420</td>	361	TATCAAGCTGGTACACGGCGACACACTTCCACTAGTGTGAGTGATCCGTATATC	420
QY <td>421</td> <td>TCTTCTCGAACCATTAACAGACTAGTATTAATTTGATCATTTGATCGTTATTTCTTGA</td> <td>480</td>	421	TCTTCTCGAACCATTAACAGACTAGTATTAATTTGATCATTTGATCGTTATTTCTTGA	480
QY <td>421</td> <td>TCTTCTCGAACCATTAACAGACTAGTATTAATTTGATCATTTGATCGTTATTTCTTGA</td> <td>480</td>	421	TCTTCTCGAACCATTAACAGACTAGTATTAATTTGATCATTTGATCGTTATTTCTTGA	480
QY <td>481</td> <td>AGCGGTTTCATTTTTTTTTTACAGACGCTTTTTTTTGAAGGTCGACATCCATTAATGCGGC</td> <td>540</td>	481	AGCGGTTTCATTTTTTTTTTACAGACGCTTTTTTTTGAAGGTCGACATCCATTAATGCGGC	540
QY <td>481</td> <td>AGCGGTTTCATTTTTTTTTTACAGACGCTTTTTTTTGAAGGTCGACATCCATTAATGCGGC</td> <td>540</td>	481	AGCGGTTTCATTTTTTTTTTACAGACGCTTTTTTTTGAAGGTCGACATCCATTAATGCGGC	540
QY <td>541</td> <td>ATAGGTGTTTACATGCGGTAATCACTTAACGTTAACACACTTTTGAAGTAATGGCTGTAAT</td> <td>600</td>	541	ATAGGTGTTTACATGCGGTAATCACTTAACGTTAACACACTTTTGAAGTAATGGCTGTAAT	600
QY <td>541</td> <td>ATAGGTGTTTACATGCGGTAATCACTTAACGTTAACACACTTTTGAAGTAATGGCTGTAAT</td> <td>600</td>	541	ATAGGTGTTTACATGCGGTAATCACTTAACGTTAACACACTTTTGAAGTAATGGCTGTAAT	600
QY <td>601</td> <td>GCGGATCTCTCCGCTACACAGCACTTTTACATTAATCTTGTGCTGTTGCAAAACCACT</td> <td>660</td>	601	GCGGATCTCTCCGCTACACAGCACTTTTACATTAATCTTGTGCTGTTGCAAAACCACT	660
QY <td>601</td> <td>GCGGATCTCTCCGCTACACAGCACTTTTACATTAATCTTGTGCTGTTGCAAAACCACT</td> <td>660</td>	601	GCGGATCTCTCCGCTACACAGCACTTTTACATTAATCTTGTGCTGTTGCAAAACCACT	660
QY <td>661</td> <td>GTTACGAATGACATCTACTGCTGTTCTGCTGACTTTATTTTTTTATTAAGTAAATACC</td> <td>720</td>	661	GTTACGAATGACATCTACTGCTGTTCTGCTGACTTTATTTTTTTATTAAGTAAATACC	720
QY <td>661</td> <td>GTTACGAATGACATCTACTGCTGTTCTGCTGACTTTATTTTTTTATTAAGTAAATACC</td> <td>720</td>	661	GTTACGAATGACATCTACTGCTGTTCTGCTGACTTTATTTTTTTATTAAGTAAATACC	720
QY <td>721</td> <td>ATTAATATGGAACAACAACCCCTGCCCTTCACTACCGGTGCGAGCGACCGCAAGATGGGG</td> <td>780</td>	721	ATTAATATGGAACAACAACCCCTGCCCTTCACTACCGGTGCGAGCGACCGCAAGATGGGG	780
QY <td>721</td> <td>ATTAATATGGAACAACAACCCCTGCCCTTCACTACCGGTGCGAGCGACCGCAAGATGGGG</td> <td>780</td>	721	ATTAATATGGAACAACAACCCCTGCCCTTCACTACCGGTGCGAGCGACCGCAAGATGGGG	780
QY <td>781</td> <td>TTCAACACGAGTGGCGACACAGATGCAACGACCTTCCAAAGCAATATCTGAGGCGGAC</td> <td>840</td>	781	TTCAACACGAGTGGCGACACAGATGCAACGACCTTCCAAAGCAATATCTGAGGCGGAC	840
QY <td>781</td> <td>TTCAACACGAGTGGCGACACAGATGCAACGACCTTCCAAAGCAATATCTGAGGCGGAC</td> <td>840</td>	781	TTCAACACGAGTGGCGACACAGATGCAACGACCTTCCAAAGCAATATCTGAGGCGGAC	840
QY <td>841</td> <td>GACGACGTAGGAGGGGTGGCCATTAACGACGGTGGCGGACATCACTTGTTCTTCCCTT</td> <td>900</td>	841	GACGACGTAGGAGGGGTGGCCATTAACGACGGTGGCGGACATCACTTGTTCTTCCCTT	900
QY <td>841</td> <td>GACGACGTAGGAGGGGTGGCCATTAACGACGGTGGCGGACATCACTTGTTCTTCCCTT</td> <td>900</td>	841	GACGACGTAGGAGGGGTGGCCATTAACGACGGTGGCGGACATCACTTGTTCTTCCCTT	900
QY <td>901</td> <td>TCTCTGTCTTCAACTTGGCCCGGACGTCTGCTAAGACCAAGGGGAATGCTGTGAGAGGA</td> <td>960</td>	901	TCTCTGTCTTCAACTTGGCCCGGACGTCTGCTAAGACCAAGGGGAATGCTGTGAGAGGA	960
QY <td>901</td> <td>TCTCTGTCTTCAACTTGGCCCGGACGTCTGCTAAGACCAAGGGGAATGCTGTGAGAGGA</td> <td>960</td>	901	TCTCTGTCTTCAACTTGGCCCGGACGTCTGCTAAGACCAAGGGGAATGCTGTGAGAGGA	960
QY <td>961</td> <td>GGTGCGGGGGCCGATTTTTTATGAGCTGGGCGAGGACAGCTGTGGCGAAACCGATCCGA</td> <td>1020</td>	961	GGTGCGGGGGCCGATTTTTTATGAGCTGGGCGAGGACAGCTGTGGCGAAACCGATCCGA	1020
QY <td>961</td> <td>GGTGCGGGGGCCGATTTTTTATGAGCTGGGCGAGGACAGCTGTGGCGAAACCGATCCGA</td> <td>1020</td>	961	GGTGCGGGGGCCGATTTTTTATGAGCTGGGCGAGGACAGCTGTGGCGAAACCGATCCGA	1020
QY <td>1021</td> <td>GCTTGCAGCAATTAACGAAGAACAGTGGGGCGCTCGCGCTACGACCGACGAGAGC</td> <td>1080</td>	1021	GCTTGCAGCAATTAACGAAGAACAGTGGGGCGCTCGCGCTACGACCGACGAGAGC	1080
QY <td>1021</td> <td>GCTTGCAGCAATTAACGAAGAACAGTGGGGCGCTCGCGCTACGACCGACGAGAGC</td> <td>1080</td>	1021	GCTTGCAGCAATTAACGAAGAACAGTGGGGCGCTCGCGCTACGACCGACGAGAGC	1080
QY <td>1081</td> <td>GGGGGTTTGTGGAGCCCTTAAGGTGGGGAAGGGGACGACCCGCTAAGGGGGGCGCATGCTC</td> <td>1140</td>	1081	GGGGGTTTGTGGAGCCCTTAAGGTGGGGAAGGGGACGACCCGCTAAGGGGGGCGCATGCTC	1140
QY <td>1081</td> <td>GGGGGTTTGTGGAGCCCTTAAGGTGGGGAAGGGGACGACCCGCTAAGGGGGGCGCATGCTC</td> <td>1140</td>	1081	GGGGGTTTGTGGAGCCCTTAAGGTGGGGAAGGGGACGACCCGCTAAGGGGGGCGCATGCTC	1140
QY <td>1141</td> <td>CAGCGCCACAGAGAAAAAAGAAAGAAAGGAGGCGCGAGATGATG</td> <td>1183</td>	1141	CAGCGCCACAGAGAAAAAAGAAAGAAAGGAGGCGCGAGATGATG	1183
QY <td>1141</td> <td>CAGCGCCACAGAGAAAAAAGAAAGAAAGGAGGCGCGAGATGATG</td> <td>1183</td>	1141	CAGCGCCACAGAGAAAAAAGAAAGAAAGGAGGCGCGAGATGATG	1183

ACCESSION M35831
VERSION M35831.1 GI:342627
KEYWORDS ribosomal protein S11; rps11 gene; rpoA gene.
SOURCE Chloroplast Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 721)
REFERENCE Markmann-Mulisch,U. and Sudbramantian,A.R.
AUTHORS Nucleotide sequence of maize chloroplast rps11 with conserved amino acid sequence between eukaryotes, bacteria and plastids
TITLE Biochem. Int. 17 (4), 655-664 (1988)
JOURNAL MEDLINE 89193772
PUBMED 3149198
COMMENT Original source text: Maize chloroplast DNA.
FEATURES Location/Qualifiers
SOURCE 1..721

```

CDS
200..631
/note="ribosomal protein S11 (rps11)"
/codon_start=1
/protein_id="AA04494.1"
/db_xref="gi:552742"
/translation="MTKAIPIKTSRKRYRIGLRNRPSPILRSKARRITGKIHYOASF
NNTIITYDPOGRVYVPSASCTGCFSGSPKSAIPACGRTVDARTYGLQRAEVMVAG
AGSGRDALRIAKSGVRLSCIRDTVTMPNNGCPFKKRL"
695..5721
CDS

```

```

/note="TPCA gene product."
/codon_start=1
/protein_id="AA084495.1"
/db_xref="GI:142629"
/translation="MWRBEITGS"

```

ORIGIN	Query Match	23.6%	Score 279;	DB 8;	Length 721;
	Best Local Similarity	100.0%	Pred. No. 1e-150;		
	Matches 279;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	408	GGATCCCTGTATCTCTTCCGACCATACAGACTAGTATTATTGATTCATGTGATCGTT	467		
DB	721	GGATCCCTGTATCTCTTCCGACCATACAGACTAGTATTATTGATTCATGTGATCGTT	662		
QY	468	TATTTCTCTTGAAAGCGGTTTCATTTTTTTTACAGAGCTTTTTTTTGGAGGTCGACA	527		
DB	661	TATTTCTCTTGAAAGCGGTTTCATTTTTTTTACAGAGCTTTTTTTTGGAGGTCGACA	602		
QY	528	TCCATTATGCGGCATAGGTGTTTACATCGCGGTATACACTTAACCGTACACCACTTTTAGC	587		
DB	601	TCCATTATGCGGCATAGGTGTTTACATCGCGGTATACACTTAACCGTACACCACTTTTAGC	542		
QY	588	AATGGCTGTATGCGGATCTCTTCGCGTACCGACACTTTTACCAATCTTGCTGCG	647		
DB	541	AATGGCTGTATGCGGATCTCTTCGCGTACCGACACTTTTACCAATCTTGCTGCG	482		
QY	648	TTGCAAACCCACTGTACGATAGCATCTACTACTGCTTTCT	686		
DB	481	TTGCAAACCCACTGTACGATAGCATCTACTACTGCTTTCT	443		

RESULT 3			
CHZMRPOA/c			
LOCUS	CHZMRPOA	1627 bp	DNA linear PLN 29-MAR-2001
DEFINITION	Zea mays chloroplast rpoa gene for RNA polymerase alpha-subunit.		
ACCESSION	X07810		
VERSION	X07810.1 GI:12466		
KEYWORDS	RNA polymerase alpha-subunit; rpoa gene; rpsl1 gene.		
SOURCE	chloroplast Zea mays		
ORGANISM	Zea mays		

REFERENCE	5	McLaughlin, W.E. and Larrina, I.M.
AUTHORS		The sequence of the maize plastid encoded rpl 22 locus
TITLE		Nucleic Acids Res. 15 (10), 4356 (1987)
JOURNAL		87231016
MEDLINE		3295778
PUBMED		
REFERENCE	6	Rodermel, S., Orlin, P. and Bogorad, L.
AUTHORS		The transcription termination region between two
TITLE		convergently-transcribed photo-regulated operons in the maize
JOURNAL		plastid chromosome contains rps14, trnK (UCU) and a putative trnfm
MEDLINE		pseudogene
PUBMED		Nucleic Acids Res. 15 (13), 5493 (1987)
REFERENCE	7	87260027
AUTHORS		3601681
TITLE		
JOURNAL		Giese, K., Subramanian, A.R., Larrina, I.M. and Bogorad, L.
MEDLINE		Nucleotide sequence, promoter analysis, and linkage mapping of the
PUBMED		unusually organized operon encoding ribosomal proteins S7 and S12
REFERENCE		in maize chloroplast
AUTHORS		J. Biol. Chem. 265 (31), 15251-15255 (1987)
TITLE		88033114
JOURNAL		2822717
MEDLINE		
PUBMED		
REFERENCE	8	Schantz, R. and Bogorad, L.
AUTHORS		Maize chloroplast genes ndh, ndhE, and psac. Sequences,
TITLE		transcripts and transcript pools
JOURNAL		Plant Mol. Biol. 11, 239-247 (1988)
MEDLINE		
PUBMED		
REFERENCE	9	Lukens, J.H. and Bogorad, L.
AUTHORS		Nucleotide sequence containing the maize chloroplast proline (UGG)
TITLE		and tryptophan (CCA) tRNA genes
JOURNAL		Nucleic Acids Res. 16 (11), 5192 (1988)
MEDLINE		88262508
PUBMED		3387224
REFERENCE	10	Ruf, M. and Kossel, H.
AUTHORS		Structure and expression of the gene coding for the alpha-subunit
TITLE		of DNA-dependent RNA polymerase from the chloroplast genome of Zea
JOURNAL		mays
MEDLINE		Nucleic Acids Res. 16 (13), 5741-5754 (1988)
PUBMED		88289331
REFERENCE		3399379
AUTHORS	11	Bowman, C.M., Barker, R.F. and Dyer, T.A.
TITLE		In wheat ctDNA, segments of ribosomal protein genes are dispersed
JOURNAL		repeats, probably conserved by nonreciprocal recombination
MEDLINE		Curr. Genet. 14 (2), 127-136 (1988)
PUBMED		89028843
REFERENCE		3180271
AUTHORS	12	Steinmüller, K., Ley, A.C., Steimetz, A.A., Sayre, R.T. and Bogorad, L.
TITLE		Characterization of the ndhC-pabG-ORF157/159 operon of maize
JOURNAL		plastid DNA and of the cyanobacterium Synechocystis sp. PCC6803
MEDLINE		Mol. Gen. Genet. 216 (1), 60-69 (1989)
PUBMED		89281491
REFERENCE		2498764
AUTHORS	13	Igloi, G.L., Meinke, A., Dory, I. and Kossel, H.
TITLE		Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon:
JOURNAL		comparison between the derived protein primary structures from
MEDLINE		various organisms with respect to functional domains
PUBMED		Mol. Gen. Genet. 221 (3), 379-394 (1990)
REFERENCE		90340289
AUTHORS		2381419
TITLE	14	Fitzky, B. and Subramanian, A.R.
JOURNAL		Nucleotide sequence and map positions of the duplicated gene for
MEDLINE		chloroplast ribosomal protein S15 in Zea mays (maize)
PUBMED		Nucleic Acids Res. 18 (11), 3407 (1990)
REFERENCE		90287750
AUTHORS		2129550
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		

QY	505	CGTCTTTTAAAGGTCGACATTCATTATTCGGCGCATAGGTGTTACATCGCGTATACAA	564
Y	445	TATATTGATCATGTAATCGTTATATTTCTCTGGAAGCGGTTCAATTTTTTTCAGA	504
Db	77139	TATATTGATCATGTAATCGTTATATTTCTCTGGAAGCGGTTCAATTTTTTTCAGA	77255
Y	505	CGTCTTTTAAAGGTCGACATTCATTATTCGGCGCATAGGTGTTACATCGCGTATACAA	564

Db	77259	CGTCTTTTGTAGAGGTGCACATCCATTATGGCGCATAGTGTATCATCGCGTATACAA	77318	
Qy	565	CTTACCGTACACCACTTTTATGCAATGCTGTAATCGGCATCTCTCCGCTACACGA	624	gene
Db	77319	CTTACCGTACACCACTTTTATGCAATGCTGTAATCGGCATCTCTCCGCTACACGA	77378	CDS
Qy	625	CCCTTTACCACTTCTGTGCTGTGCAACCCACCTGATGAGATATAGTGGCTGT	684	
Db	77379	CCCTTTACCACTTCTGTGCTGTGCAACCCACCTGATGAGATATAGTGGCTGT	77438	
Qy	685	CT 686		
Db	77439	CT 77440		
RESULT 5				
AE009947		141182 bp	DNA	circular
LOCUS	AE009947	141182 bp	DNA	circular
DEFINITION	Saccharum hybrid cultivar SP-80-3280 chloroplast, complete genome.			
ACCESSION	AE009947			
VERSION	AE009947.2	GI:50080438		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	1 (bases 1 to 141182)			
TITLE	Calza, T. Jr., Carraro, D.M., Benatti, M.R., Barbosa, A.C.,			
JOURNAL	Kitaajima, J.P., and Carter, H.			
REFERENCE	Structural features and transcript editing analysis of sugarcane			
AUTHORS	(Saccharum officinarum L.) chloroplast genome			
JOURNAL	Unpublished			
TITLE	2 (bases 1 to 141182)			
AUTHORS	Calza, T. Jr.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (21-JAN-2002) Ciencias Biologicas, Escola Superior de			
AUTHORS	Agricultura Luiz de Queiroz - Universidade de Sao Paulo, Av. Padua			
JOURNAL	Dias, 11 - Agronomia, Piracicaba, SP 13418-900, Brazil			
COMMENT	On Jul 9, 2004 this sequence version replaced gi:48478610.			
FEATURES				
source	location/Qualifiers			
	1..141182			
	/organism="Saccharum hybrid cultivar SP-80-3280"			
	/organism="Saccharum hybrid cultivar SP-80-3280"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:193079"			
	complement(1139..510)			
gene	/gene="rpl14"			
	/locus_tag="PS001"			
	complement(139..510)			
CDS	/gene="rpl14"			
	/locus_tag="PS001"			
	/codon_start=1			
	/transl_table=11			
	/product="ribosomal protein L14"			
	/protein_id="AA14631.1"			
	/db_xref="GI:48478611"			
	/translation="MIQPTLLVADNSGARKLNCIRVIGAAGNORVARIQDVIIVAVI			
	KDAVQWPLERSEVIRAVIVTRKEFKDDDIIRYDNNAAVVIDQKNPSTVFGA			
	VAEEDRELNTKIVSLAPEVL"			
	complement(620..2109)			
gene	/gene="rpl16"			
	/locus_tag="PS002"			
	complement(join(620..1021,2101..2109))			
CDS	/gene="rpl16"			
	/locus_tag="PS002"			
	/codon_start=1			
	/transl_table=11			
	/product="ribosomal protein L16"			
	/protein_id="AA14632.1"			
	/db_xref="GI:48478612"			
	/translation="WLSPKRTTRFKQHGKRNAGKSGRNHICFGKVALQVLEPAMITA			
	ROI EAGRAMTRVARRGKIWVRI PPDKEVTIRPTETRMGSGKSGSPSEVVAVVKPGR I			
	LYEMSGSETVARRAISAISAKMPIRSGLRIE I"			
	complement(2254..2928)			
gene	/gene="rps3"			
	/locus_tag="PS003"			
CDS	complement(2254..2928)			
	/gene="rps3"			
	/locus_tag="PS003"			
	/codon_start=1			
	/transl_table=11			
	/product="ribosomal protein S3"			
	/protein_id="AA14633.1"			
	/db_xref="GI:48478613"			
	/translation="MGQINPLGFRLLGTTQNHHSFWPAQPKVYSEGLQEDKIRNCIK			
	NYIOKRRKSGNRKMSDSSSEVITHTIEIOKEIDITHVTHIGFPLNLKKKGAIELE			
	KDLOKEVSNVQRLNIAIEKVEPRQPNILAEVIAFQKRPVSPKAMKAIELTK			
	ADIKGIQIAGLAKETARACIKRGRPLQTIIRAKIDYCCYPIRTIYGLGVKVI			
	IFVEEE"			
	complement(2988..3434)			
gene	/gene="rpl22"			
	/locus_tag="PS004"			
CDS	complement(2988..3434)			
	/gene="rpl22"			
	/locus_tag="PS004"			
	/codon_start=1			
	/transl_table=11			
	/product="ribosomal protein L22"			
	/protein_id="AA14634.1"			
	/db_xref="GI:48478614"			
	/translation="MTSFKLVKTPRIKKKKGRKLARKVPTDRLLKEEFVFAQKRI			
	HMSVFAQRLVDIEIRMYEETMIIMLPYASYPILKLVSAANAATYRQEDKTN			
	LFTLKAVERSTMTKKPRPARGRSYSIKKTWCNITITLVINVKSK"			
	complement(3527..3808)			
gene	/gene="rps19"			
	/locus_tag="PS005"			
CDS	complement(3527..3808)			
	/gene="rps19"			
	/locus_tag="PS005"			
	/codon_start=1			
	/transl_table=11			
	/product="ribosomal protein S19"			
	/protein_id="AA14635.1"			
	/db_xref="GI:48478615"			
	/translation="MTRKTNPFVARRLAKIEKVNKEKEEIIIVTSRASSILPAMV			
	GHTIATNGKEHTPIYTNNMVGKGEFPTPHFISVESTROTYSRR"			
	complement(3828..4052)			
gene	/locus_tag="PS007"			
	complement(3828..4052)			
CDS	/locus_tag="PS007"			
	/note="orf75"			
	/codon_start=1			
	/transl_table=11			
	/product="hypothetical protein 75"			
	/protein_id="AA14636.1"			
	/db_xref="GI:48478616"			
	/translation="MENCIRGNCNNMAGSERBELNPRMVDSSTALIHATSAPYPAKG			
	PSLSSTHYISGISDLHTSIRIVDIGCHSLK"			
	3939..4013			
gene	/gene="trnh"			
	/locus_tag="PS006"			
tRNA	3939..4013			
	/gene="trnh"			
	/locus_tag="PS006"			
	/product="trna-His"			
	4175..4588			
gene	/locus_tag="PS009"			
	4175..4588			
CDS	/locus_tag="PS009"			
	/note="orf137"			
	/codon_start=1			
	/transl_table=11			
	/product="hypothetical protein 137"			
	/protein_id="AA14637.1"			

[illegible]

Every Match

20.5%; Score 242; DB 8; Length 141182;

Best Local Similarity	100.0%	Pred. No. 9.1e-129	Mismatches	0	Indels	0	Gaps	0
Matches 242; Conservative	0							
QY	445	TATTATTGATCATGATCGTTTATTTCTCTTGAAGCGGTTTCATTTTTTTTACAGA	504					
Db	139485	TATTATTGATCATGATCGTTTATTTCTCTTGAAGCGGTTTCATTTTTTTTACAGA	139544					
QY	505	CGGCTTTTTTTTGGAGGTGAGATCATTTATGCGGATGGTGTACATCGCGATACAA	564					
Db	139545	CGGCTTTTTTTTGGAGGTGAGATCATTTATGCGGATGGTGTACATCGCGATACAA	139604					
QY	565	CTTAACCGTAGACACACTTTTACCAATGGCTCGTATATGGCGATCTCTCCGCTACGCA	624					
Db	139605	CTTAACCGTAGACACACTTTTACCAATGGCTCGTATATGGCGATCTCTCCGCTACGCA	139664					
QY	625	CGTTTATACCAATACTTGTGCTCGTTGCAAAACCACTGTACGAATAGCATCTTGTCTTT	684					
Db	139665	CGTTTATACCAATACTTGTGCTCGTTGCAAAACCACTGTACGAATAGCATCTTGTCTTT	139724					
QY	685	CT 686						
Db	139725	CT 139726						
RESULT 6								
LOCUS	AP006714	141182 bp	DNA	circular	PLN 06-JUL-2004			
DEFINITION	Saccharum officinarum chloroplast DNA, complete genome.							
ACCESSION	AP006714							
VERSION	AP006714.1	GI:49659489						
KEYWORDS								
SOURCE								
ORGANISM	chloroplast Saccharum officinarum							
REFERENCE	1	Aano, T., Tauduku, T., Takahashi, S., Shimada, H. and Kadowaki, K.						
AUTHORS	Complete Nucleotide Sequence of the Sugarcane (Saccharum							
TITLE	officinarum) Chloroplast Genome: A Comparative Analysis of Four							
JOURNAL	Monocot Chloroplast Genomes							
REFERENCE	DNA Res. 11, 93-99 (2004)							
AUTHORS	2 (bases 1 to 141182)							
TITLE	Kadowaki, K.							
JOURNAL	Direct Submission							
REFERENCE	Submitted (17-NOV-2003) Koh-ichi Kadowaki, National Institute of							
AUTHORS	Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki							
TITLE	305-6602, Japan (E-mail: kadowaki@affrc.go.jp, Tel: 81-29-838-7449,							
JOURNAL	Fax: 81-29-838-7408)							
FEATURES								
source	Location/Qualifiers							
	1..141182							
	/organism="Saccharum officinarum"							
	/organelle="plastid:chloroplast"							
	/mol_type="genomic DNA"							
	/db_xref="taxon:4547"							
	/note="Saccharum officinarum L."							
	complement(join(93000..93028,93569..93800,69929..70042))							
	/codon_start=1							
	/rname="table=11							
	/product="ribosomal protein S12"							
	/protein_id="BAD27348.1"							
	/db_xref="GI:49659567"							
	/translation="MPYKQLIRNAROPIRNARSALKGPORGTCAVYTIINPK							
	PNSLRKRVRLVLTSGFEITAVIPGIGHNIOEHSVLVVRGRVXKDLPGVYRIIRGTL							
	DAVAYNRQGSKTKGAKKPK"							
	join(complement(93569..93800),complement(93000..93028),							
	complement(69929..70042),130431..130662,131231)							
	/gene="rps12"							
	join(complement(69929..70042),130431..130662,							
	131203..131231)							
	/gene="rps12"							
	/codon_start=1							

gene
CDS

transl_table=1
product="ribosomal protein S12"
protein_id="BAD27316.1"
db_xref="GI:49659535"
translation="MPTVKQLINRARPPIRNARKSALKGCPQRGTCAVYTNPKPNSLRVRARLRLTSGEITAYIPGIGHNLQHSVLVVRGAVVDLPGVRIIRFGLTDAVAIKRRQGRSKYGAKKPEK"
complement(join(1389..1423,3902..3938))
/gene="trnk"
complement(join(1389..1423,3902..3938))
/gene="crnk"
product="rRNA-Lys"
note="codon recognized: AAA"
complement(1673..3274)
/gene="matk"
complement(1673..3274)
/gene="matk"
/codon_start=1
/transl_table=1
product="maturase"
protein_id="BAD27271.1"
db_xref="GI:49659490"
translation="MYHMLINREMLPSLIQVEIQMKPEGYSKKORHPFYPLIFOEIYAPHADYGLANGSEBEVEIGCNKKKFSSILKYKLIIIMYQNPLSNWYNPNDRLEFDHNYYSSFFSYQSILSBGFIVAEIPLSGQUSCEBKEIPKQMLQSIHSIFPLBDKPFHLHYLSHIEIPYIHLEILVOLLEVRIQVPSELHREFLAHYSWMNLSITSMKSIFLKKENORLHFELNYSVSEYEFLFLFKROSSCLRLTSSGTELEBIIFSQMEHGWAYPOEFPKRITWFEDPPLMHYRVXGKAIIASKGTLKKKMSYLVINFSQFPYFWQOVRRLNOLNTSCDFPLGYLSSVPIINTLVNRNMLENSPLIDTRMKKNTPVATPLVSLSKAOPCTGSGHPISKPVMWTDLSDPDILDRGRICRNI.FHHSSSKKOTLYRLKLTIRYCSARTLARHKSTVTRPMORLSVFLEBFTEBEQVFSLMFTKHIFSFBGSOSERIVLYDIIRINDLVNPLTN"
complement(join(4503..4720,5555..5594))
/gene="tps16"
complement(join(4503..4720,5555..5594))
/gene="tps16"
/codon_start=1
/transl_table=1
product="ribosomal protein S16"
protein_id="BAD27272.1"
db_xref="GI:49659491"
translation="MWKRLKRCGRKQAITYRIAIVDRSRREGDRLRKGYTPICKQPLNVPAIIFYLEGAQPTRVYDILRKAFEPDKERTLS"
complement(7124..7195)
/gene="lrnq"
complement(7124..7195)
/gene="lrnq"
product="rRNA-Gln"
note="codon recognized: CAA"
7550..7735
/gene="psbk"
7550..7735
/gene="psbk"

gene
CDS

codon_start=1
transl_table=1
product="PSII K-protein"
protein_id="BAD27273.1"
db_xref="GI:49659492"
translation="WPNILSLTICIFNSVLCPTSFPPAKLDEATAINFPIYDVNPVLFVLFPLAFVMAAVSFR"
8124..8234
/gene="psbi"
8124..8234
/gene="psbt"
/codon_start=1
/transl_table=1
product="PSII I-protein"
protein_id="BAD27274.1"
db_xref="GI:49659493"
translation="MTLTKLPHYTYVIFPVSLFIIFGLSNDGRNPGREE"
complement(8361..8448)
/gene="trns"
complement(8361..8448)

gene
tRNA

transl_table=1
product="ribosomal protein S12"
protein_id="BAD27316.1"
db_xref="GI:49659535"
translation="MPTVKQLINRARPPIRNARKSALKGCPQRGTCAVYTNPKPNSLRVRARLRLTSGEITAYIPGIGHNLQHSVLVVRGAVVDLPGVRIIRFGLTDAVAIKRRQGRSKYGAKKPEK"
complement(join(1389..1423,3902..3938))
/gene="trnk"
complement(join(1389..1423,3902..3938))
/gene="crnk"
product="rRNA-Lys"
note="codon recognized: AAA"
complement(1673..3274)
/gene="matk"
complement(1673..3274)
/gene="matk"
/codon_start=1
/transl_table=1
product="maturase"
protein_id="BAD27271.1"
db_xref="GI:49659490"
translation="MYHMLINREMLPSLIQVEIQMKPEGYSKKORHPFYPLIFOEIYAPHADYGLANGSEBEVEIGCNKKKFSSILKYKLIIIMYQNPLSNWYNPNDRLEFDHNYYSSFFSYQSILSBGFIVAEIPLSGQUSCEBKEIPKQMLQSIHSIFPLBDKPFHLHYLSHIEIPYIHLEILVOLLEVRIQVPSELHREFLAHYSWMNLSITSMKSIFLKKENORLHFELNYSVSEYEFLFLFKROSSCLRLTSSGTELEBIIFSQMEHGWAYPOEFPKRITWFEDPPLMHYRVXGKAIIASKGTLKKKMSYLVINFSQFPYFWQOVRRLNOLNTSCDFPLGYLSSVPIINTLVNRNMLENSPLIDTRMKKNTPVATPLVSLSKAOPCTGSGHPISKPVMWTDLSDPDILDRGRICRNI.FHHSSSKKOTLYRLKLTIRYCSARTLARHKSTVTRPMORLSVFLEBFTEBEQVFSLMFTKHIFSFBGSOSERIVLYDIIRINDLVNPLTN"
complement(join(4503..4720,5555..5594))
/gene="tps16"
complement(join(4503..4720,5555..5594))
/gene="tps16"
/codon_start=1
/transl_table=1
product="ribosomal protein S16"
protein_id="BAD27272.1"
db_xref="GI:49659491"
translation="MWKRLKRCGRKQAITYRIAIVDRSRREGDRLRKGYTPICKQPLNVPAIIFYLEGAQPTRVYDILRKAFEPDKERTLS"
complement(7124..7195)
/gene="lrnq"
complement(7124..7195)
/gene="lrnq"
product="rRNA-Gln"
note="codon recognized: CAA"
7550..7735
/gene="psbk"
7550..7735
/gene="psbk"

gene
CDS

codon_start=1
transl_table=1
product="PSII K-protein"
protein_id="BAD27273.1"
db_xref="GI:49659492"
translation="WPNILSLTICIFNSVLCPTSFPPAKLDEATAINFPIYDVNPVLFVLFPLAFVMAAVSFR"
8124..8234
/gene="psbi"
8124..8234
/gene="psbt"
/codon_start=1
/transl_table=1
product="PSII I-protein"
protein_id="BAD27274.1"
db_xref="GI:49659493"
translation="MTLTKLPHYTYVIFPVSLFIIFGLSNDGRNPGREE"
complement(8361..8448)
/gene="trns"
complement(8361..8448)

gene
tRNA

transl_table=1
product="ribosomal protein S12"
protein_id="BAD27316.1"
db_xref="GI:49659535"
translation="MPTVKQLINRARPPIRNARKSALKGCPQRGTCAVYTNPKPNSLRVRARLRLTSGEITAYIPGIGHNLQHSVLVVRGAVVDLPGVRIIRFGLTDAVAIKRRQGRSKYGAKKPEK"
complement(join(1389..1423,3902..3938))
/gene="trnk"
complement(join(1389..1423,3902..3938))
/gene="crnk"
product="rRNA-Lys"
note="codon recognized: AAA"
complement(1673..3274)
/gene="matk"
complement(1673..3274)
/gene="matk"
/codon_start=1
/transl_table=1
product="maturase"
protein_id="BAD27271.1"
db_xref="GI:49659490"
translation="MYHMLINREMLPSLIQVEIQMKPEGYSKKORHPFYPLIFOEIYAPHADYGLANGSEBEVEIGCNKKKFSSILKYKLIIIMYQNPLSNWYNPNDRLEFDHNYYSSFFSYQSILSBGFIVAEIPLSGQUSCEBKEIPKQMLQSIHSIFPLBDKPFHLHYLSHIEIPYIHLEILVOLLEVRIQVPSELHREFLAHYSWMNLSITSMKSIFLKKENORLHFELNYSVSEYEFLFLFKROSSCLRLTSSGTELEBIIFSQMEHGWAYPOEFPKRITWFEDPPLMHYRVXGKAIIASKGTLKKKMSYLVINFSQFPYFWQOVRRLNOLNTSCDFPLGYLSSVPIINTLVNRNMLENSPLIDTRMKKNTPVATPLVSLSKAOPCTGSGHPISKPVMWTDLSDPDILDRGRICRNI.FHHSSSKKOTLYRLKLTIRYCSARTLARHKSTVTRPMORLSVFLEBFTEBEQVFSLMFTKHIFSFBGSOSERIVLYDIIRINDLVNPLTN"
complement(join(4503..4720,5555..5594))
/gene="tps16"
complement(join(4503..4720,5555..5594))
/gene="tps16"
/codon_start=1
/transl_table=1
product="ribosomal protein S16"
protein_id="BAD27272.1"
db_xref="GI:49659491"
translation="MWKRLKRCGRKQAITYRIAIVDRSRREGDRLRKGYTPICKQPLNVPAIIFYLEGAQPTRVYDILRKAFEPDKERTLS"
complement(7124..7195)
/gene="lrnq"
complement(7124..7195)
/gene="lrnq"
product="rRNA-Gln"
note="codon recognized: CAA"
7550..7735
/gene="psbk"
7550..7735
/gene="psbk"

gene /gene="trnS"
/product="tRNA-Ser"
/note="codon recognized: AGC"
9014. .9145
/gene="orf43"
9014. .9145
/gene="orf43"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAD27275.1"
/db_xref="GI:49659494"
/translation="MNKGFOKSIPLFIPLFFPRILFLPHPCNRERMRKESFLKK"
9423. .10486
/gene="psbD"
9423. .10486
/gene="psbD"
/codon_start=1
/transl_table=1
/product="PSII D2 protein"
/protein_id="BAD27276.1"
/db_xref="GI:49659495"
/translation="MTIAGVGRVTKENDLFDIMDWLRDRFVFWGSGLLLPCCAYF
ALGSGFTTQTSMTWTHGLASSYLBGCFNLTAVSTPANSLSHLLMLGPEAGQDF
TRWQIGGLMTFVALHGFALIGFMARQFELARVQLRPNATISGSDIIVPVSFLI
YPLGSGFEPASFGVALIFRFILFFGFSHWTLNLPFHMNGVAGVCLALCAIHGHT
VENTLFEEDGANTFRFANFPQAEITYSMVTANFWSQITGVAFSNRKLHFLFMLFVP
VTHGLMSIIGVGLALNLRAYDFVSOEIRAADEPPEFTYKNILLNEGIRAWMAAD
OPHENLIPREVLRGNAL"
10434. .11855
/gene="psbC"
10434. .11855
/gene="psbC"
/codon_start=1
/transl_table=1
/product="PSII 44kD protein"
/protein_id="BAD27277.1"
/db_xref="GI:49659496"
/translation="MKILYSLRPFYHWETLFGTFVLAGRDOETTGFAWNGARLILN
LSGLKIGHVAHAGLIFVMAAGANMLFEVAHFVPKPKMYEQGLILPHLATLGWGVGGF
GCVLTPFPVPSGVLHLSASVAGFGIYHALGPELLEESFPFGVMDRNMKTI
IGTHILIGLGAFLILYKALYFGGVDTWARGGDKRTNLTLSPGVIRGYLLSKSP
GEGHIVSYDDLEDILIGHWMLGSICTUGCTIMHILTRPAMARAAPWMSGAVLSL
MALSVFGIACCPWFNNNTAYPSEFTGPTGEASQAQFTFLVRDQKLGAVSQAQRP
TOLGYLMKSPGGEVIFGGETFRFMDLRAFWLELRGNGLDLSRLKDDIQPWERSS
AYTMHAYGSLNSVSGVATEINAVNVSPRSMLATSHFVIGFPFVGHLMHGRARA
AAAGPEKGIIDRILEPVLNMTPLN"
12915. .12985
/gene="trnS"
12915. .12985
/gene="trnS"
/product="tRNA-Ser"
/note="codon recognized: UCA"
12441. .12629
/gene="psbZ"
12441. .12629
/gene="psbZ"
/codon_start=1
/transl_table=1
/product="PSII Z-protein"
/protein_id="BAD27278.1"
/db_xref="GI:49659497"
/translation="MTIAPOLAVFALLIATSSVLISVPLVPASPDGSSNNKNVFSCT
SLMIGLVPLVATLNSLIS"
12915. .12985
/gene="trnG"
12915. .12985
/gene="trnG"
/product="tRNA-Gly"
/note="codon recognized: GGC"
complement(13477. .13550)
/gene="trnH"
complement(13477. .13550)

QY 163 TTTTATATATGTGAAAAGAGTAATTTCACTTTGGGCCACTTTTATTACGATATTTT 222
| | | | |
Db 920 TTTTATATATGTGAAAAGAGTAATTTCACTTTGGGCCACTTTTATTACGATATTTT 861
| | | | |
QY 223 ACTTTATACACCTTTTAACTGATGTTTCACTTTTGACGAGGTAATCTTACCTTTGTTT 282
| | | | |
Db 860 ACTTTATACACCTTTTAACTGATGTTTCACTTTTGACGAGGTAATCTTACCTTTGTTT 801
| | | | |
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 333
| | | | |
Db 800 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 750
| | | | |
RESULT 10
AY452735/c 4895 bp DNA circular SYN 19-DEC-2003
AY452735/c
LOCUS Reporter vector pActXN, complete sequence.
DEFINITION AY452735
ACCESSION AY452735
VERSION AY452735.1 GI:39636979
KEYWORDS
SOURCE Reporter vector pActXN
ORGANISM Reporter vector pActXN
other sequences: artificial sequences; vectors.
REFERENCE 1 (bases 1 to 4895)
AUTHORS Vickers,C.E., Xue,G.P. and Greshoff,P.M.
TITLE A synthetic xylanase as a novel reporter in plants
JOURNAL Plant Cell Rep. 22 (2), 135-140 (2003)
MEDLINE 22867549
PUBMED 12845475
REFERENCE 2 (bases 1 to 4895)
AUTHORS Vickers,C.E.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2003) ARC Centre for Integrative Legume Research,
The University of Queensland, Room 213, John Hines Building (69),
St. Lucia, QLD 4072, Australia
FEATURES
source
1. .4895
/organism="Reporter vector pActXN"
/mol_type="other DNA"
/db_xref="taxon:255329"
267..2274
/note="sequence from cloning vector pBR322"
468..1328
/note="ampicillin resistance protein; bla"
/codon_start=1
/product="beta-lactamase"
/protein_id="AAR29085.1"
/db_xref="GI:39636980"
/translacion="MSIQHFRVALIPFPAAFCFLVFAHPETLVKVKDAEDQLGARVGY
IETDLSGKILRESFRRPFPMSTFVKLCGAVLSRIDGQEQLGRVHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAIIETMSDNTANLLITIGSKELTAFLHMGDVTPL
DRWPELNEAIPIRDERDTTPVMAATTLRKILTGSELLLAGROOLIDMEADKVAGPL
LRSLAPGMFLADSKSGAGERSRGIITALLPGDKPRSLIVITYTGSQAIDMERRRQLA
EIGASLIKHW"
2345..2367
/note="T7 RNA polymerase promoter"
2426..3338
/note="from rice actin gene; Act1"
3339..3792
/note="intron 1 from rice actin gene; Act1"
3800..4513
/gene="xyna"
/gene="xyna"
/gene="xyna"
/note="synthetic xylanase"
/codon_start=1
/product="xylanase"
/protein_id="AAR29086.1"
/db_xref="GI:39636981"
/translacion="MASNGKFTVGNQNHQNVNDGSEYELIDNTGNSMTLGS
ATPKAEMAAVNRGNFLARGLDPSOKKADYDYLIDYATVYKOTASNSLGV
YGMFQNRGLNVPVLEYITIEDMDVWVVDAGCKMTTIDGAYKIFQMDHTPTNGGS
ETPKYFVSVRQDKRTSGHITVSDHFKEMAKQGMIGNLVEVALNAEGWSSGVADVTL

LDVYTPKSGSPATSAPR"
4535..4872
/note="from rbcS gene"
4895
/note="SP6 RNA polymerase transcription initiation site"
ORIGIN
Query Match 14.5%; Score 171; DB 12; Length 4895;
Best Local Similarity 100.0%; Pred.No.1.9e-87;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 163 TTTTATATATGTGAAAAGAGTAATTTCACTTTGGGCCACTTTTATTACGATATTTT 222
| | | | |
Db 2593 TTTTATATATGTGAAAAGAGTAATTTCACTTTGGGCCACTTTTATTACGATATTTT 2534
| | | | |
QY 223 ACTTTATACACCTTTTAACTGATGTTTCACTTTTGACGAGGTAATCTTACCTTTGTTT 282
| | | | |
Db 2533 ACTTTATACACCTTTTAACTGATGTTTCACTTTTGACGAGGTAATCTTACCTTTGTTT 2474
| | | | |
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 333
| | | | |
Db 2473 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 2423
| | | | |
RESULT 11
AX384394/c 9359 bp DNA linear PAT 19-MAR-2002
AX384394/c
LOCUS Sequence 3 from Patent WO0214524.
DEFINITION AX384394
ACCESSION AX384394
VERSION AX384394.1 GI:19577678
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences: artificial sequences.
REFERENCE 1
AUTHORS van der Valk,P., van Dun,C.M., Smeekens,S.C. and Proveniers,M.C.
TITLE Inhibition of generative propagation in genetically modified
JOURNAL herbicide resistant grasses
Patent: WO 0214524-A 3 21-FEB-2002;
Advanta Seeds B.V. (NL)
FEATURES
source
1. .9359
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
839..1699
/note="unnamed protein product; Beta-lactamase gene
(Ampr)"
/codon_start=1
/translacion="1
/protein_id="CAD28571.1"
/db_xref="GI:19577679"
/translacion="MSIQHFRVALIPFPAAFCFLVFAHPETLVKVKDAEDQLGARVGY
IETDLSGKILRESFRRPFPMSTFVKLCGAVLSRIDGQEQLGRVHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAIIETMSDNTANLLITIGSKELTAFLHMGDVTPL
DRWPELNEAIPIRDERDTTPVMAATTLRKILTGSELLLAGROOLIDMEADKVAGPL
LRSLAPGMFLADSKSGAGERSRGIITALLPGDKPRSLIVITYTGSQAIDMERRRQLA
EIGASLIKHW"
2941..4920
/note="ubi-promoter from maize"
4921..6400
/note="Ath1 gene from Arabidopsis thaliana"
6401..6672
/note="poly-A signal from the nopaline synthetase gene
from Agrobacteriu m tumefaciens"
7434..8084
/note="first exon-intron combination from Ubi-maize"
8085..9119
/note="Hygromycin resistance gene from Escherichia coli"
9120..9359
ORIGIN
polyA_signal
Query Match 14.5%; Score 171; DB 6; Length 9359;


```
ORGANISM      synthetic construct
REFERENCE      other sequences; artificial sequences.
1
AUTHORS        Bakker,H.A., Florack,D.E. and Bosch,H.J.
TITLE          GntIIi (udp-n-acetylglucosamine:beta-d-mannoside
JOURNAL        beta(1,4)-n-acetylglucosaminyltransferase ii) expression in plants
                Patent: WO 03078614-A 10 25-SEP-2003;
                Plant Research International B.V. (NL)
FEATURES
SOURCE          Location/Qualifiers
                1..7545
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Synthetic"
ORIGIN
Query Match      14.3%; Score 169; DB 6; Length 7545;
Best Local Similarity 100.0%; Pred. No. 3e-86;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 163 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACGATATTT 222
    |||||
DB 1473 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACGATATTT 1414
QY 223 ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 282
    |||||
DB 1413 ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 1354
QY 283 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 331
    |||||
DB 1353 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 1305

RESULT 15
AX840289/c     11643 bp   DNA   linear   PAT 16-DEC-2003
LOCUS          AX840289
DEFINITION     Sequence 11 from Patent WO03078614.
ACCESSION      AX840289
VERSION        AX840289.1 GI:39978688
KEYWORDS
SOURCE         synthetic construct
ORGANISM       synthetic construct
OTHER SEQUENCES; artificial sequences.
1
REFERENCE      1
AUTHORS        Bakker,H.A., Florack,D.E. and Bosch,H.J.
TITLE          GntIIi (udp-n-acetylglucosamine:beta-d-mannoside
JOURNAL        beta(1,4)-n-acetylglucosaminyltransferase ii) expression in plants
                Patent: WO 03078614-A 11 25-SEP-2003;
                Plant Research International B.V. (NL)
FEATURES
SOURCE          Location/Qualifiers
                1..11643
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Synthetic"
ORIGIN
Query Match      14.3%; Score 169; DB 6; Length 11643;
Best Local Similarity 100.0%; Pred. No. 3.2e-86;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 163 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACGATATTT 222
    |||||
DB 5573 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACGATATTT 5514
QY 223 ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 282
    |||||
DB 5513 ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 5454
QY 283 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 331
    |||||
DB 5453 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 5405
```

Search completed: February 10, 2005, 13:13:42
Job time : 7299.97 secs


```

XX PI Behr CF, Hironaka C, Heck GR, You J;
XX DR WPI; 2002-165871/22.
XX PT Novel DNA construct useful for producing a corn plant that tolerates
XX application of glyphosate herbicide, comprises two transgene expression
XX cassettes.
XX PS Claim 3; Page 16; 25pp; English.
XX CC The invention relates to a DNA construct (I) comprising a first (F1) and
XX a second (F2) expression cassette. F1 of the DNA construct, in operable
XX linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
XX peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
XX phosphate synthase (EPSPS) gene (III), and transcriptional terminator
XX (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,
XX (III), and (IV). Also included are two DNA sequences from corn plant
XX PV-ZMGT32(nk603) of 498 or 1183bp, or primers and probes derived from
XX them (used to detect transgene junction points). The construct is useful
XX for producing a corn plant that tolerates application of glyphosate
XX herbicide by transforming a corn cell with, selecting the corn cell for
XX tolerance to application of glyphosate, and growing the corn cell into a
XX fertile corn plant. The present sequence is the 498bp amplicon consisting
XX of corn genomic sequence, vector sequences and Rice Act1 promoter
XX sequences
XX SQ Sequence 498 BP; 148 A; 114 C; 131 G; 105 T; 0 U; 0 Other;

Query Match 100.0%; Score 498; DB 6; Length 498;
Best Local Similarity 100.0%; Pred. No. 9.6e-246;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGATCCAAATCGGAGCTGAATGTTGGAAGAAAGAGACAGAGACCTCAGCTT 60
DB 1 AATGATCCAAATCGGAGCTGAATGTTGGAAGAAAGAGACAGAGACCTCAGCTT 60
QY 61 TCCAGGGTGAAGTATCAGAGGATTTACCGCCATGCTTTTATGAGACAAAGAGGGAG 120
DB 61 TCCAGGGTGAAGTATCAGAGGATTTACCGCCATGCTTTTATGAGACAAAGAGGGAG 120
QY 121 GAGGTAAACAGATGAGCATCGCGCTGGAAGTTTCTGCAAGATGCGGAAGCTGTTCC 180
DB 121 GAGGTAAACAGATGAGCATCGCGCTGGAAGTTTCTGCAAGATGCGGAAGCTGTTCC 180
QY 121 GAGGTAAACAGATGAGCATCGCGCTGGAAGTTTCTGCAAGATGCGGAAGCTGTTCC 180
DB 121 GAGGTAAACAGATGAGCATCGCGCTGGAAGTTTCTGCAAGATGCGGAAGCTGTTCC 180
QY 181 AGCGCGGTGGCATTTGGGCGAGACTCCTCTCTGGGCGATGAGCGATCTTTCTGTG 240
DB 181 AGCGCGGTGGCATTTGGGCGAGACTCCTCTCTGGGCGATGAGCGATCTTTCTGTG 240
QY 241 GCATTTCCAAACCTAGAGACGTGCTCCTGTGGGTGCTCGGCGACCAAGCTTTAG 300
DB 241 GCATTTCCAAACCTAGAGACGTGCTCCTGTGGGTGCTCGGCGACCAAGCTTTAG 300
QY 241 GCATTTCCAAACCTAGAGACGTGCTCCTGTGGGTGCTCGGCGACCAAGCTTTAG 300
DB 241 GCATTTCCAAACCTAGAGACGTGCTCCTGTGGGTGCTCGGCGACCAAGCTTTAG 300
QY 301 CGGCGCCAGCGGTGATCCATGATATCCCTTAGGGGCGCGGCTTAACAGCTTACTC 360
DB 301 CGGCGCCAGCGGTGATCCATGATATCCCTTAGGGGCGCGGCTTAACAGCTTACTC 360
QY 301 CGGCGCCAGCGGTGATCCATGATATCCCTTAGGGGCGCGGCTTAACAGCTTACTC 360
DB 301 CGGCGCCAGCGGTGATCCATGATATCCCTTAGGGGCGCGGCTTAACAGCTTACTC 360
QY 361 GAGGTCAATTCATGCTTGAAGAGAGTGGGATAGCCAAATTAAGAAAGAGTAAAG 420
DB 361 GAGGTCAATTCATGCTTGAAGAGAGTGGGATAGCCAAATTAAGAAAGAGTAAAG 420
QY 421 TTACCGGTCAAAAGTGAACATCAGTTAAAGGTGTTAAAGTAAATATCGGTAAATA 480
DB 421 TTACCGGTCAAAAGTGAACATCAGTTAAAGGTGTTAAAGTAAATATCGGTAAATA 480
QY 421 TTACCGGTCAAAAGTGAACATCAGTTAAAGGTGTTAAAGTAAATATCGGTAAATA 480
DB 421 TTACCGGTCAAAAGTGAACATCAGTTAAAGGTGTTAAAGTAAATATCGGTAAATA 480
QY 481 AAGGTGCGCCAAAGTGAA 498
DB 481 AAGGTGCGCCAAAGTGAA 498

```

RESULT 2
ABK15245/c
ID ABK15245 standard; DNA; 1183 BP.

```

XX AC ABK15245;
XX DT 08-MAY-2002 (first entry)
XX DE Corn nk603 3' transgene/genomic sequence PCR fragment.
XX KW ds; rice actin 1 promoter; RA1; RA1 intron;
XX KW chloroplast transit peptide gene; glyphosate resistance; corn;
XX KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;
XX KW transcriptional terminator; cauliflower mosaic virus 35S promoter;
XX KW PV-ZMGT32; transgenic; nk603.
XX OS Zea mays.
XX OS Escherichia coli.
XX OS Agrobacterium tumefaciens.
XX OS Synthetic.
XX OS Chimeric.
XX FH Key
XX FT misc_feature
XX FT 1..164
XX FT /tag= a
XX FT /note= "T-AGRTU.nos vector sequence"
XX FT misc_feature
XX FT 165..381
XX FT /tag= b
XX FT /note= "Vector sequence"
XX FT misc_feature
XX FT 382..686
XX FT /tag= c
XX FT /note= "Corn plasmid genes rps11/rpoA partial sequence"
XX FT misc_feature
XX FT 687..1183
XX FT /tag= d
XX FT /note= "Corn genomic sequence"
XX PN EP167531-A1.
XX XX 15-JUN-2001; 2001EB-00202314.
XX PD 02-JAN-2002.
XX PF 22-JUN-2000; 2000US-0213567P.
XX PR 13-OCT-2000; 2000US-0240014P.
XX PR 13-OCT-2000; 2000US-0241215P.
XX PA (MONSANTO TECHNOLOGY LLC.
XX PI Behr CF, Hironaka C, Heck GR, You J;
XX DR WPI; 2002-165871/22.
XX PT Novel DNA construct useful for producing a corn plant that tolerates
XX application of glyphosate herbicide, comprises two transgene expression
XX cassettes.
XX PS Claim 3; Page 16-17; 25pp; English.
XX CC The invention relates to a DNA construct (I) comprising a first (F1) and
XX a second (F2) expression cassette. F1 of the DNA construct, in operable
XX linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
XX peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
XX phosphate synthase (EPSPS) gene (III), and transcriptional terminator
XX (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,
XX (III), and (IV). Also included are two DNA sequences from corn plant
XX PV-ZMGT32(nk603) of 498 or 1183bp, or primers and probes derived from
XX them (used to detect transgene junction points). The construct is useful
XX for producing a corn plant that tolerates application of glyphosate
XX herbicide by transforming a corn cell with, selecting the corn cell for
XX tolerance to application of glyphosate, and growing the corn cell into a
XX fertile corn plant. The present sequence is the 1183bp amplicon
XX consisting of the Agrobacterium transcriptional terminator sequence,
XX vector sequences and corn sequences
XX SQ Sequence 1183 BP; 295 A; 289 C; 280 G; 319 T; 0 U; 0 Other;

```


PT used to produce transgenic plants e.g. banana, wheat, maize or rice,
 PT having resistance or tolerance to glyphosate herbicide.
 XX
 PS Claim 19, Page 55-56; 87pp; English.

CC The invention relates to rice 5-enolpyruvylshikimate phosphate synthase
 CC (EPSPS) genomic DNA (AAC87188). The invention also relates to an
 CC expression cassette comprising, in the 5'-3' direction, one or more
 CC transcriptional enhancer elements selected from AAC87190-C87196), the
 CC rice EPSPS promoter, genomic DNA encoding a rice EPSPS chloroplast
 CC transit peptide, genomic DNA encoding a EPSPS protein modified such that
 CC it is resistant to glyphosate (AAC87189), and a transcriptional
 CC terminator. The glyphosate resistant EPSPS contains a region (AAB29793)
 CC containing two amino acid substitutions relative to the corresponding
 CC wild-type region (AAB29792). The invention also encompasses plant genomic
 CC EPSPS sequences identified via screening with a rice EPSPS intronic
 CC sequence; vectors and host plant cells comprising a nucleic acid sequence
 CC of the invention; transgenic plants (and tissues and seeds thereof)
 CC comprising a nucleic acid sequence of the invention, optionally further
 CC transformed with a DNA encoding an insect, fungal, viral, bacterial,
 CC nematode, stress or herbicide resistance protein; and methods of
 CC producing the transgenic plants of the invention. The nucleic acids and
 CC constructs of the invention are used to produce a wide variety of
 CC morphologically normal, glyphosate resistant plants. The glyphosate
 CC resistant plants produced are particularly maize, soybean, cotton,
 CC sugarcane and canola, but also other field crops, fruits and vegetables,
 CC turf and forage grasses and nut-producing plants. The plants are
 CC optionally resistant to insects, fungi, viruses, bacteria, nematodes,
 CC stress, desiccation and/or other herbicides. They can be used in the
 CC production of a herbicidal target for the high throughput in vitro
 CC screening of potential herbicides. The present sequence represents an
 CC enhancer element which may be used in the rice EPSPS expression cassette
 CC of the invention

SQ Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;

Query Match 14.1%; Score 70; DB 3; Length 1501;

Best Local Similarity 100.0%; Pred. No. 1.8e-25;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 415

DB 750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 809

QY 416 TAAGATTACC 425

DB 810 TAAGATTACC 819

RESULT 5

ID AAC89325 standard; DNA; 1501 BP.

XX AAC89325;

DT 07-MAR-2001 (first entry)

DE Rice actin enhancer.

XX 5-enolpyruvylshikimate phosphate synthase; EPSPS; herbicide resistance;

XX glyphosate; ds.

OS Oryza sp.

PN MO200066747-A1.

PD 09-NOV-2000.

PF 20-APR-2000; 2000WO-GB001572.

PR 29-APR-1999; 99GB-00009967.

PR 29-APR-1999; 99GB-00009969.

PR 29-APR-1999; 99GB-00009972.

PR 29-APR-1999; 99GB-00009981.

PR 29-JUL-1999; 99GB-00017835.

PR 29-JUL-1999; 99GB-00017836.

PR 29-JUL-1999; 99GB-00017843.

PR 21-DEC-1999; 99GB-00030202.

PR 21-DEC-1999; 99GB-00030210.

PR 21-DEC-1999; 99GB-00030212.

XX (ZENEC) ZENEC LTD.

PI Hawkes TR, Warner SAJ, Andrews CJ, Bachoo S, Pickerill AP;

XX WPI; 2000-679764/66.

XX Example 20; Page 54; 98pp; English.

CC The present invention relates to an Oryza sp. 5-enolpyruvylshikimate

CC phosphate synthase (EPSPS) gene. Vectors containing the gene may be used

CC to produce plant tissues and fertile whole plants which are substantially

CC tolerant or substantially resistant to glyphosate herbicide and to

CC produce a herbicidal target which is used for high throughput in vitro

CC screening of potential herbicides

SQ Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;

Query Match 14.1%; Score 70; DB 3; Length 1501;

Best Local Similarity 100.0%; Pred. No. 1.8e-25;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 415

DB 750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 809

QY 416 TAAGATTACC 425

DB 810 TAAGATTACC 819

RESULT 6

ID ADC84543 standard; DNA; 2181 BP.

XX ADC84543;

DT 01-JAN-2004 (first entry)

DE Rice actin promoter encoding sequence.

XX expression cassette; ds.

XX Melanomy caliginosus.

OS MO2003027257-A2.

PN 03-APR-2003.

PD 27-SEP-2002; 2002WO-US030895.

PF 27-SEP-2001; 2001US-0325607P.

PR (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Levin JZ, Dietrich R, Budziszewski GJ;

XX WPI; 2003-354651/33.

XX New isolated nucleic acid molecule encoding a polypeptide with a 3'-5'

XX exonuclease domain, useful in molecular biology and transformation, in

XX particular for reproducing and predictably manipulating gene expression

PT in a plant cell.
 XX
 XX
 PS Disclosure; Page 106-107; 108bp; English.
 XX
 CC The present invention relates to an invention that results in the
 CC expression of the endogenous nucleotide sequence in plant cell being
 CC increased. In this method the plant cell comprises a second expression
 CC cassette comprising a nucleic acid molecule, where the expression of the
 CC nucleic acid molecule of interest in the plant cell is decreased as
 CC compared to expression of the nucleic acid molecule of interest in a
 CC plant cell lacking the first expression cassette. The methods and
 CC compositions of the present invention are useful in molecular biology and
 CC transformation, in particular for reproducing and predictably
 CC manipulating gene expression in a plant cell. The present sequence is a
 CC nucleotide sequence of rice actin promoter.
 XX
 SO Sequence 2181 BP; 602 A; 518 C; 430 G; 631 T; 0 U; 0 Other;
 Query Match 14.1%; Score 70; DB 10; Length 2181;
 Best Local Similarity 100.0%; Pred. No. 1.8e-25;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 TACTCGAGTCATTCATATGCTTGAAGAAGAGTCGGGATAGTCCAAATTAACAAGG 415
 DB 755 TACTCGAGTCATTCATATGCTTGAAGAAGAGTCGGGATAGTCCAAATTAACAAGG 814
 QY 416 TAAGATTACC 425
 DB 815 TAAGATTACC 824
 RESULT 7
 AAT80052
 ID AAT80052 standard; cDNA; 2199 BP.
 XX
 AC AAT80052;
 XX
 DT 25-MAR-2003 (revised)
 DT 04-NOV-1997 (first entry)
 XX
 DE Rice actin 1 gene promoter region.
 XX
 KW Promoter; rice; actin 1 gene; cytoplasm; cytoskeleton; extension growth;
 KW cytoplasmic streaming; cell division; Rac1; maize; Adh1 promoter; ss.
 XX
 OS Oryza sativa.
 XX
 PN US5641876-A.
 XX
 PD 24-JUN-1997.
 XX
 PF 27-OCT-1993; 93US-00144602.
 XX
 PR 05-JAN-1990; 90US-00461490.
 PR 18-SEP-1991; 91US-00762680.
 XX
 PA (CORR) CORNELL RES FOUND INC.
 XX
 FI Wu R, Mcelroy D;
 XX
 DR WPI; 1997-340996/31.
 XX
 PT Nucleic acid containing the promoter of the rice actin-1 gene - used to
 PT direct efficient expression of foreign genes in rice.
 XX
 PS Claim 3; Col 33-36; 29pp; English.
 XX
 CC This sequence represents the promoter region from the rice actin-1 (Rac1)
 CC gene. Cytoplasmic actin is a fundamental and essential component of the
 CC eukaryotic cell and cytoskeleton. In higher plant cells, a number of
 CC cellular processes, such as cytoplasmic streaming, extension growth and
 CC cell division are thought to involve the cytoskeletal actin protein. All
 CC of the studied plant actins consist of four exons of conserved length,

CC separated by 3 introns of variable length. In rice there are at least 8
 CC actin-like sequences per haploid genome. Rac1 encodes a transcript that
 CC is relatively abundant in all rice tissues. This sequence is an example
 CC of a nucleic acid molecule of the invention. The nucleic acid molecules
 CC of the invention contain the Rac1 gene (or fragments) with promoter
 CC activity in monocotyledonous plants. The promoter is used to direct
 CC expression of foreign genes in transgenic rice and other plants. The
 CC actin promoter is more efficient in rice transformation than previously
 CC proposed promoters (e.g. 5 times more active than the maize Adh1
 CC promoter) and has constitutive activity in space and time. (Updated on 25
 CC -MAR-2003 to correct PF field.)
 XX
 SO Sequence 2199 BP; 603 A; 515 C; 446 G; 635 T; 0 U; 0 Other;
 Query Match 14.1%; Score 70; DB 2; Length 2199;
 Best Local Similarity 100.0%; Pred. No. 1.8e-25;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 TACTCGAGTCATTCATATGCTTGAAGAAGAGTCGGGATAGTCCAAATTAACAAGG 415
 DB 806 TACTCGAGTCATTCATATGCTTGAAGAAGAGTCGGGATAGTCCAAATTAACAAGG 865
 QY 416 TAAGATTACC 425
 DB 866 TAAGATTACC 875
 RESULT 8
 AAQ12707
 ID AAQ12707 standard; DNA; 5510 BP.
 XX
 AC AAQ12707;
 XX
 DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 27-SEP-1991 (first entry)
 XX
 DE Rice actin gene Rac1.
 XX
 KW Promoter; transgenic plant; monocotyledonous; ss.
 XX
 OS Oryza sativa; variety IR26.
 XX
 FH Key
 FT Location/Qualifiers
 FT 1..2071
 FT /tag= a
 FT 1027..1038
 FT /tag= b
 FT 1078..1088
 FT /tag= c
 FT /rpt_type= direct
 FT 1301..1333
 FT /tag= d
 FT /rpt_type= tandem (imperfect)
 FT 1352..1386
 FT /tag= e
 FT /rpt_type= tandem (imperfect)
 FT /note= "eight pentamers"
 FT 1465..1505
 FT /tag= f
 FT /label= purine-rich region
 FT 1609..1617
 FT /tag= g
 FT 1647..1653
 FT /tag= h
 FT /label= transcription initiation signal
 FT 1650..3741
 FT /tag= p
 FT /label= genomic clone pRac1
 FT 1650..1728
 FT /tag= i
 FT /note= "transcription initiation site"

```

FT repeat_region 1650. .1676
FT /tag= j
FT /rpl_type= tandem
FT /note= "GC-rich - repeated (A/T)CC triplets"
FT misc_RNA 1728. .1731
FT /tag= k
FT /label= 5' acceptor splice site
FT intron 1729. .2043
FT /tag= 1
FT /label= 5' acceptor splice site
FT misc_RNA 2008. .2012
FT /tag= m
FT /label= putative branch point splice site
FT misc_RNA 2034. .2044
FT /tag= n
FT /label= 3' donor splice site
FT CDS 2051. .3600
FT /tag= o
FT /label= 3' donor splice site

XX PN WO9109948-A.
XX PD 11-JUL-1991.
XX PF 05-JAN-1990; 90US-00461490.
XX PR 05-JAN-1990; 90US-00461490.
XX PA (CORR ) CORNELL RES FOUND INC.
XX PI Wu R, Mcelroy D;
XX DR WPI; 1991-222901/30.
XX PT Rice actin gene promoter - used to confer high level expression of
XX PT foreign genes in transgenic rice and other plants.
XX PS Claim 8; Page 10; 47pp; English.
XX CC The seugence was obtd. from a clone isolated from a rice genomic library
XX CC screened with a heterologous actin probe. The CDS encodes a 377 amino
XX CC acid actin protein of 41.9 kD. The gene is interrupted by three introns
XX CC (position not disclosed). Analysis also revealed an additional intron 5'
XX CC of the translation initiation region. This 5' intron separates a 79 bp GC
XX CC -rich 5' noncoding exon from the exon coding the translation initiation
XX CC codon. This is one of few cases of a 5' noncoding exon in a plant gene.
XX CC The promoter region can be used to confer high level expression of
XX CC foreign genes in transgenic rice and other agronomically important
XX CC plants. It is five times more active than the maize Adh1 promoter in
XX CC trans-formed rice. (Updated on 25-MAR-2003 to correct PA field.)
XX CC (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 5510 BP; 1466 A; 1262 C; 1163 G; 1619 T; 0 U; 0 Other;
XX

Query Match 14.1%; Score 70; DB 2; Length 5510;
Best Local Similarity 100.0%; Pred. No. 1.7e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATTAACAAGG 415
DB 809 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATTAACAAGG 868
QY 416 TAAGATTACC 425
DB 869 TAAGATTACC 878

```

```

XX DE Rice actin 1 gene.
XX DE
XX KW Promoter; rice; actin 1 gene; cytoplasm; cytoskeleton; extension growth;
XX KW cytoplasmic streaming; cell division; RAC1; maize; Adh1 promoter; ss.
XX OS Oryza. sativa.
XX PN US5641876-A.
XX PD 24-JUN-1997.
XX PF 27-OCT-1993; 93US-00144602.
XX PR 05-JAN-1990; 90US-00461490.
XX PR 18-SEP-1991; 91US-00762680.
XX PA (CORR ) CORNELL RES FOUND INC.
XX PI Wu R, Mcelroy D;
XX DR WPI; 1997-340996/31.
XX PF Nucleic acid containing the promoter of the rice actin-1 gene - used to
XX PT direct efficient expression of foreign genes in rice.
XX PS Example 1; Col 3-7; 29pp; English.
XX CC This sequence represents a the rice actin-1 (Rac1) gene amplified using
XX CC AAT80054. Cytoplasmic actin is a fundamental and essential component of
XX CC the eukaryotic cell and cytoskeleton. In higher plant cells, a number of
XX CC cellular processes, such as cytoplasmic streaming, extension growth and
XX CC cell division are thought to involve the cytoskeletal actin protein. All
XX CC of the studied plant actins consist of four exons of conserved length,
XX CC separated by 3 introns of variable length. In rice there are at least 8
XX CC actin-like sequences per haploid genome. Rac1 encodes a transcript that
XX CC is relatively abundant in all rice tissues. This sequence is an example
XX CC of a nucleic acid molecule of the invention. The nucleic acid molecules
XX CC of the invention contain the Rac1 gene, or fragments with promoter
XX CC activity (see AAT80052 and AAT80053) in monocotyledonous plants. The
XX CC promoter is used to direct expression of foreign genes in transgenic rice
XX CC and other plants. The actin promoter is more efficient in rice
XX CC transformation than previously proposed promoters (e.g. 5 times more
XX CC active than the maize Adh1 promoter) and has constitutive activity in
XX CC space and time. (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 5643 BP; 1500 A; 1283 C; 1200 G; 1660 T; 0 U; 0 Other;
XX

Query Match 14.1%; Score 70; DB 2; Length 5643;
Best Local Similarity 100.0%; Pred. No. 1.7e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATTAACAAGG 415
DB 809 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATTAACAAGG 868
QY 416 TAAGATTACC 425
DB 869 TAAGATTACC 878

```

```

RESULT 10
ID ABK11039 standard; DNA; 9359 BP.
AC ABK11039;
XX
XX 05-JUN-2002 (first entry)
XX
XX pVDH636 vector containing Arabidopsis Ath1 gene.
XX
XX pVDH636 vector; circular; cyclic; grass; plant; herbicide resistance;
XX baseball; cricket; football; golf; rugby; soccer; tennis; lawn; park;

```

KW athletic field; animal feedstuff; grass flowering inhibitor; Ath1; gene;
 KM ds.
 XX Arabidopsis thaliana.
 OS Zea mays.
 OS Agrobacterium tumefaciens.
 OS Oryza sativa.
 OS Escherichia coli.
 OS Cauliflower mosaic virus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 4930..6360
 FT /*tag= a
 FT /product= "Ath1"
 FT /partial
 FT /note= "No stop codon"
 PN WO200214524-A2.
 XX
 PD 21-FEB-2002.
 PF 16-AUG-2001; 2001WO-EP009572.
 XX
 PR 18-AUG-2000; 2000US-0226422P.
 PR 27-NOV-2000; 2000US-0253274P.
 PR 27-NOV-2000; 2000US-0253327P.
 PR 22-JUN-2001; 2001US-0300220P.
 XX
 PA (ADVA-) ADVANTA SEEDS BV.
 PI Van Der Valk P, Van Dun CMP, Smeekens SCM, Proveniers MCG;
 XX WPI: 2002-257611/30.
 DR P-PSDB; AAU76514.
 XX
 PT New genetically modified grasses that express inhibited generative
 PT propagation, or herbicide resistance, useful for forage (e.g. cattle
 PT feedstuff) or amenity purposes (e.g. for use in an athletic field, lawn
 PT or park).
 XX
 PS Example 1; Fig 6; 56pp; English.
 XX
 CC The invention relates to a grass plant, which has been genetically
 CC modified to substantially inhibit generative propagation and carry
 CC herbicide resistance. The grass is useful for growth and/or propagation
 CC of grasses. The grass is used to play at least one sport (e.g. baseball,
 CC cricket, football, golf, rugby, soccer or tennis), or used at least in a
 CC portion of an athletic field, lawn or park. The grass is fed to animal
 CC (e.g. cattle, goat, horse or sheep) or used as an animal feedstuff. The
 CC present sequence represents the coding sequence of pVDH636 vector
 CC containing the Arabidopsis Ath1 gene, which was used to inhibit flowering
 CC in grasses
 XX
 SQ Sequence 9359 BP; 2377 A; 2222 C; 2133 G; 2624 T; 0 U; 3 Other;
 XX
 Query Match 14.1%; Score 70; DB 6; Length 9359;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 TACTCGAGGTCAATTCATATGCTTGAGAGAGAGTCCGAGATAGTCCAAATAAACAAGG 415
 DB 6688 TACTCGAGGTCAATTCATATGCTTGAGAGAGAGTCCGAGATAGTCCAAATAAACAAGG 6747
 QY 416 TAAAGTTACC 425
 DB 6748 TAAAGTTACC 6757
 RESULT 11
 ID ABK10687 standard; DNA; 9359 BP.
 XX

AC ABK10687;
 XX 05-JUN-2002 (first entry)
 XX
 DE Transformation vector plasmid pVDH636 DNA.
 XX
 KW Grass; plant; transgenic; flowering inhibition; inflorescence; gene; ds;
 KW tiller production; delayed heading; gibberellic acid; phytohormone; golf;
 KW genetically modified grass; athletic field; sport; baseball; cricket;
 KW football; rugby; soccer; tennis; lawn; landscaping; cattle; horse; sheep;
 KW goat; animal feedstuff; Arabidopsis thaliana; Ath1; homeobox gene;
 KW circular; cyclic; pVDH636.
 XX
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 4930..6351
 FT /*tag= a
 FT /product= "Arabidopsis thaliana Ath1"
 FT
 PN WO200214486-A2.
 XX
 PD 21-FEB-2002.
 PF 16-AUG-2001; 2001WO-EP009570.
 XX
 PR 18-AUG-2000; 2000US-0226422P.
 PR 27-NOV-2000; 2000US-0253274P.
 PR 27-NOV-2000; 2000US-0253327P.
 PR 22-JUN-2001; 2001US-0300220P.
 XX
 PA (ADVA-) ADVANTA SEEDS BV.
 PI Van Der Valk P, Van Dun CMP, Smeekens SCM, Proveniers MCG;
 XX WPI: 2002-257603/30.
 DR P-PSDB; AAU76885.
 XX
 PT New genetically modified grass useful for growing and/or propagating
 PT grass in athletic fields (for sports such as baseball, cricket, football,
 PT golf, rugby, soccer and tennis), or as animal feedstuff for cattle, goat,
 PT horse and sheep.
 XX
 PS Example 1; Fig 6; 45pp; English.
 XX
 CC The invention relates to a grass plant which has been genetically
 CC modified to substantially inhibit generative propagation. The genetic
 CC modification may result in a heritable change in one or more plant
 CC characteristics such as inhibition of flowering (or substantial delay
 CC that amounts to inhibition), absence of inflorescence, increased
 CC production of tillers, delayed heading and inhibition of the
 CC developmental switch from vegetative to generative growth. A method of
 CC making a grass involves transforming the grass with a nucleic acid which
 CC interferes with metabolism of gibberellic acid. A grass can be created by
 CC applying a phytohormone to at least partially relieve or reverse a change
 CC in plant characteristic resulting from genetic modification. The
 CC genetically modified grass is useful for growing and/or propagating grass
 CC in athletic fields (for sports such as baseball, cricket, football, golf,
 CC rugby, soccer and tennis), lawns, parks and other types of landscaping.
 CC The grass is also useful as an animal feedstuff for cattle, goats, horses
 CC and sheep, due to its increased vegetative growth, improved digestibility
 CC and/or nutritional value as animal feedstuff. This sequence represents
 CC plasmid pVDH636 DNA. This plasmid was used as a transformation vector for
 CC production of transgenic grasses expressing the Arabidopsis thaliana
 CC homeobox gene, Ath1
 XX
 SQ Sequence 9359 BP; 2377 A; 2222 C; 2133 G; 2624 T; 0 U; 3 Other;
 XX
 Query Match 14.1%; Score 70; DB 6; Length 9359;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 TACTCGAGGTCAATTCATATGCTTGAGAGAGAGTCCGAGATAGTCCAAATAAACAAGG 415

DB 6688 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGG 6747
QY 416 TAAGATTACC 425
DB 6748 TAAGATTACC 6757

RESULT 12
ADG69912
ID ADG69912 standard; DNA; 1259 BP.
AC ADG69912;
DT 11-MAR-2004 (first entry)
DE Actin promoter and nucleotide sequence.

XX transgenic plant; ferulic acid esterase; FAE; enzyme; phenolic acid;
KW plant; cell wall; improved digestibility; biomass conversion;
KM highly fermentable carbohydrate; gene; ds.
XX
OS Synthetic.
PN WO200268666-A1.
PD 06-SEP-2002.
PP 16-NOV-2001; 2001WO-US043588.
PR 17-NOV-2000; 2000US-0249608P.
PX (GEMV) GENENCOR INT INC.
PY
PI Dunn-Coleman N, Langdon T, Morris P;
DR WPI; 2002-698675/75.

XX New transgenic plant comprising an expression cassette with a promoter
PT operably linked to a ferulic acid esterase encoding polynucleotide,
PT useful in improving digestibility for livestock and enhancing biomass
PT conversion.

XX Disclosure; Fig 50A-B; 208pp; English.
XX
XX The present invention describes a transgenic plant (I) comprising an
CC expression cassette with a promoter operably linked to a ferulic acid
CC esterase encoding polynucleotide. Also described: (1) controlling (M1)
CC the level of phenolic acids in plant cell walls of a transgenic plant by
CC introducing into the plant an expression cassette comprising a promoter
CC operably linked to a ferulic acid esterase encoding polynucleotide; and
CC (2) a transgenic plant (II) produced by (M1). The transgenic plants are
CC useful in improving digestibility for livestock and enhancing biomass
CC conversion. The method is useful for enhancing the production of more
CC highly fermentable carbohydrates in plants, especially forage grasses.
CC The expression cassette is useful for controlling the level of phenolic
CC acids in plant cell walls of a transgenic plant. The present sequence is
CC used in the exemplification of the present invention.
XX
XX Sequence 1259 BP; 337 A; 332 C; 268 G; 322 T; 0 U; 0 Other;

Query Match 13.7%; Score 68; DB 6; Length 1259;
Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTGATGCTTGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 417
DB 16 CTCGAGTCATTGATGCTTGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 75
QY 418 AGATTACC 425
DB 76 AGATTACC 83

RESULT 13
ADN88926
ID ADN88926 standard; DNA; 1259 BP.
AC ADN88926;
DT 01-JUL-2004 (first entry)
DE Actin promoter.

XX Transgenic plant; ferulic acid esterase; fermentable carbohydrate;
KW animal feed; phenolic acid; actin; promoter; ds.
XX
OS Synthetic.
PN WO2003043411-A2.
PD 30-MAY-2003.
PP 21-MAY-2002; 2002WO-US016239.
PR 16-NOV-2001; 2001US-00991209.
PR 16-NOV-2001; 2001WO-US043588.
PX (GEMV) GENENCOR INT INC.
PY
PI Dunn-Coleman N, Langdon T, Morris P;
DR WPI; 2003-457563/43.

XX New transgenic plant comprising an expression cassette with a promoter
PT operably linked to ferulic acid esterase encoding polynucleotide, useful
PT for enhancing availability of fermentable carbohydrates as feed for grass
PT -fed animals.

XX Disclosure; Fig 50; 190pp; English.
XX
XX The present invention relates to transgenic plants comprising an
CC expression cassette having a promoter operably linked to a ferulic acid
CC esterase (FAE) encoding polynucleotide. The transgenic plants are useful
CC for enhancing the availability of fermentable carbohydrates as feed for
CC grass-fed animals. The methods are useful for reducing the levels of
CC phenolic acids in the cell walls available for cross-linking either by
CC directly disrupting ester bonds linking phenolics and lignins to cell
CC wall polysaccharides or by preventing excessive ferulic acid of cell wall
CC carbohydrates. The present sequence was used to illustrate the invention.
XX
XX Sequence 1259 BP; 337 A; 332 C; 268 G; 322 T; 0 U; 0 Other;

Query Match 13.7%; Score 68; DB 11; Length 1259;
Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTGATGCTTGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 417
DB 16 CTCGAGTCATTGATGCTTGAGAGAGTCGGATAGTCCAAATAAACAAGGTA 75
QY 418 AGATTACC 425
DB 76 AGATTACC 83

RESULT 14
AAT80053
ID AAT80053 standard; cDNA; 1392 BP.
AC AAT80053;
DT 25-MAR-2003 (revised)
DT 04-NOV-1997 (first entry)
DE Rice actin 1 gene promoter fragment.

XX Promoter; rice; actin 1 gene; cytoplasm; cytoskeleton; extension growth;
 KW cytoplasmic streaming; cell division; Rac1; maize; Adh1 promoter; ss.
 XX
 OS *Oryza sativa*.
 XX
 PN US561876-A.
 XX
 PD 24-JUN-1997.
 XX
 PF 27-OCT-1993; 93US-00144602.
 XX
 PR 05-JAN-1990; 90US-00461490.
 PR 18-SEP-1991; 91US-00762680.
 XX
 PA (CORR) CORNELL RES FOUND INC.
 XX
 PI Wu R, McElroy D;
 DR WPI; 1997-340996/31.
 XX
 PT Nucleic acid containing the promoter of the rice actin-1 gene - used to
 PT direct efficient expression of foreign genes in rice.
 XX
 PS Claim 4; Col 35-38; 29pp; English.
 XX
 CC This sequence represents a fragment from the promoter region of the rice
 CC actin-1 (Rac1) gene. Cytoplasmic actin is a fundamental and essential
 CC component of the eukaryotic cell and cytoskeleton. In higher plant cells,
 CC a number of cellular processes, such as cytoplasmic streaming, extension
 CC growth and cell division are thought to involve the cytoskeletal actin
 CC protein. All of the studied plant actins consist of four exons of
 CC conserved length, separated by 3 introns of variable length. In rice
 CC there are at least 8 actin-like sequences per haploid genome. Rac1
 CC encodes a transcript that is relatively abundant in all rice tissues.
 CC This sequence is an example of a nucleic acid molecule of the invention.
 CC The nucleic acid molecule of the invention contains the Rac1 gene (or
 CC fragments) with promoter activity in monocotyledonous plants. The
 CC promoter is used to direct expression of foreign genes in transgenic rice
 CC and other plants. The actin promoter is more efficient in rice
 CC transformation than previously proposed promoters (e.g. 5 times more
 CC active than the maize Adh1 promoter) and has constitutive activity in
 CC space and time. (Updated on 25-MAR-2003 to correct PF field.)
 CC
 SQ Sequence 1392 BP; 376 A; 343 C; 309 G; 364 T; 0 U; 0 Other;
 XX
 QY Query Match 13.7%; Score 68; DB 2; Length 1392;
 Db Best Local Similarity 100.0%; Pred. No. 1.9e-24;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 358 CTCGAGGTCAATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGGTA 417
 Db 1 CTCGAGGTCAATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGGTA 60
 QY 418 AGATTACC 425
 Db 61 AGATTACC 68
 XX
 RESULT 15
 ADG69892
 ID ADG69892 standard; DNA; 3657 BP.
 XX
 AC ADG69892;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Vector pTP8-5 nucleotide sequence.
 XX
 KM transgenic plant; ferulic acid esterase; FAE; enzyme; phenolic acid;
 KM plant; cell wall; improved digestibility; biomass conversion;
 KM highly fermentable carbohydrate; gene; ds.
 XX

OS Synthetic.
 XX
 PN WO200268666-A1.
 XX
 PD 06-SEP-2002.
 XX
 PF 16-NOV-2001; 2001WO-US043588.
 XX
 PR 17-NOV-2000; 2000US-0249608P.
 XX
 PA (GENEV) GENENCOR INT INC.
 XX
 PI Dunn-Coleman N, Langdon T, Morris P;
 DR WPI; 2002-698675/75.
 DR P-PSDB; ADG69893.
 XX
 PT New transgenic plant comprising an expression cassette with a promoter
 PT operably linked to a ferulic acid esterase encoding polynucleotide,
 PT useful in improving digestibility for livestock and enhancing biomass
 PT conversion.
 XX
 PS Disclosure; Fig 36B-E; 208pp; English.
 XX
 CC The present invention describes a transgenic plant (I) comprising an
 CC expression cassette with a promoter operably linked to a ferulic acid
 CC esterase encoding polynucleotide. Also described: (i) controlling (M1)
 CC the level of phenolic acids in plant cell walls of a transgenic plant by
 CC introducing into the plant an expression cassette comprising a promoter
 CC operably linked to a ferulic acid esterase encoding polynucleotide; and
 CC (2) a transgenic plant (II) produced by (M1). The transgenic plants are
 CC useful in improving digestibility for livestock and enhancing biomass
 CC conversion. The method is useful for enhancing the production of more
 CC highly fermentable carbohydrates in plants, especially forage grasses.
 CC The expression cassette is useful for controlling the level of phenolic
 CC acids in plant cell walls of a transgenic plant. The present sequence is
 CC used in the exemplification of the present invention.
 CC
 SQ Sequence 3657 BP; 916 A; 1012 C; 868 G; 860 T; 0 U; 1 Other;
 XX
 QY Query Match 13.7%; Score 68; DB 6; Length 3657;
 Db Best Local Similarity 100.0%; Pred. No. 1.9e-24;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 358 CTCGAGGTCAATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGGTA 417
 Db 2430 CTCGAGGTCAATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGGTA 2489
 QY 418 AGATTACC 425
 Db 2490 AGATTACC 2497
 XX
 Search completed: February 10, 2005, 13:43:58
 Job time : 516.751 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 7997.05 Seconds
(without alignments)
5630.828 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183

Sequence: 1 gacgtatctatcagatgagatgaggg.....aaggaaagcgcgagatgatg 1183

Scoring table: OLIGO-NTC

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	305	25.8	537	B2310383 1c53c07.b
2	305	25.8	553	BH883307 h22a01.b
3	305	25.8	648	B2314790 h22a01.b
4	305	25.8	680	B2625170 1g96e05.g
5	305	25.8	717	B2320382 h207h12.g
6	305	25.8	738	CC161210 1176c10.b
7	305	25.8	776	BH878576 h882h12.b
8	305	25.8	824	CD438581 EL01N0515
9	305	25.8	920	CC357906 PUHHU36TD
10	280	22.7	600	B2375835 1e62b08.g
11	275	22.2	600	B2373363 1e78f08.b
12	273	22.1	613	CC025412 3591_146
13	256	21.6	646	B2359336 1d72d02.b
14	253	21.4	650	B2372017 1e62b08.g
15	241	20.4	428	B2363755 1d90d01.g
16	238	20.1	835	B2697563 PUCCK44TD
17	238	20.1	955	B2747937 PUCSC09TD
18	195	16.5	1010	CM002222 ZMMBHg000
19	184	15.6	736	CC804949 1h38d04.g
20	183	15.5	453	B2310528 1c55b09.b
21	183	15.5	606	B2778381 1c63b07.g
22	182	15.4	613	CC156828 1g15h05.g
23	177	15.0	1047	CM001399 ZMMBHg000
24	177	15.0	1060	CL990345 ZMMBHf000

C 25	177	15.0	1062	9	CM002030	CM002030 ZMMBHg000
C 26	169	14.3	1013	6	CM002257	CM002257 ZMMBHh000
C 27	168	14.2	670	6	CB867772	CB867772 HC01D06w
C 28	166	14.0	166	8	BH415256	BH415256 1007042C0
C 29	160	13.5	672	8	B2349476	B2349476 h41b05.g
C 30	160	13.5	955	9	CL239469	CL239469 ZMMBHh0058
C 31	159	13.4	500	8	B2316096	B2316096 1a67f08.b
C 32	159	13.4	500	8	B2322609	B2322609 1a67f08.g
C 33	158	13.4	971	9	CM001948	CM001948 ZMMBHg000
C 34	152	12.8	633	9	CG159311	CG159311 PUKK31TD
C 35	145	12.3	569	8	B2341607	B2341607 1c46f03.g
C 36	145	12.3	606	8	B2329855	B2329855 hv88d04.g
C 37	145	12.3	632	8	B2332275	B2332275 hv86e12.g
C 38	145	12.3	707	8	CM006474	CM006474 ZMMBLA000
C 39	145	12.3	782	8	B2781027	B2781027 1123b10.g
C 40	145	12.3	845	9	CM009185	CM009185 ZMMBLA001
C 41	134	11.3	474	7	CV061066	CV061066 BNEI64e8
C 42	134	11.3	490	7	CV059547	CV059547 BNEI49c6
C 43	134	11.3	529	7	CV057962	CV057962 BNEI32g8
C 44	134	11.3	544	7	CV060284	CV060284 BNEI56d9
C 45	134	11.3	549	7	CV060835	CV060835 BNEI61g2

ALIGNMENTS

RESULT 1
B2310383
LOCUS
DEFINITION
B2310383 537 bp DNA linear GSS 06-NOV-2002
1c53c07.b1 WGS-ZmayrF (JM107 adapted methyl filtered) Zea mays
ACCESSION
B2310383
VERSION
B2310383.1 GI:24672266
KEYWORDS
SOURCE
ORGANISM
GSS.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 537)
Rabinowitz, P.D., O'Shaughnessy, A.L., Ballie, V., Dedhia, N.,
Katzemburger, F., King, L., Miller, B., Muller, S., Nascimben, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: 1c53 row: c column: 07
Seq primer: -21M13UnivFwd
Clas: shotgun
High quality sequence stop: 537.
Location/Qualifiers
1..537
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="1c53c07"
/lab_host="JM107 or DH5a"
/clone_lib="WGS-ZmayrF (JM107 adapted methyl filtered)"
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, .b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

TITLE
JOURNAL
COMMENT

FEATURES
source

ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.5e-157; Mismatches 0; Indels 0; Gaps 0;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ACAACCTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 441
DB 91 ACACACTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 150
QY 442 TAGATATTATGATCATTTGAATCGTTTATCTCTTGAAGCGGTTTCAATTTTTTTTAC 501
DB 151 TAGATATTATGATCATTTGAATCGTTTATCTCTTGAAGCGGTTTCAATTTTTTTTAC 210
QY 502 AGACGCTTTTTTTAGAGGTGAGCATCATTAATGCGGATAGGTGTATACATCGGATTA 561
DB 211 AGACGCTTTTTTTAGAGGTGAGCATCATTAATGCGGATAGGTGTATACATCGGATTA 270
QY 562 CAACCTTAACCGTACACCACTTTTGAAGAGCTGTGAATGCGGATCTCTTCGCGTACCA 621
DB 271 CAACCTTAACCGTACACCACTTTTGAAGAGCTGTGAATGCGGATCTCTTCGCGTACCA 330
QY 622 GCACCTTTTACCAATCTTGTCTGCTGTGCAAAACCACTGTACGAATAGCATCTACT 681
DB 331 GCACCTTTTACCAATCTTGTCTGCTGTGCAAAACCACTGTACGAATAGCATCTACT 390
QY 682 GTTCT 686
DB 391 GTTCT 395

RESULT 2
LOCUS BH883307/c 553 bp DNA linear GSS 05-AUG-2002
DEFINITION hW43f05.g1 MGS-Zmaysef (JM107 adapted methyl filtered) Zea mays
ACCESSION BH883307
VERSION BH883307.1 GI:22119204
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 553)
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,
Katzemburger, F., King, L., Miller, B., Muller, S., Nascento, L.,
Zutavern, T., McCombie, W.R., and Martensen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)

AUTHORS Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org

COMMENT Plate: hW43 row: f column: 05
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 553.
Location/Qualifiers
1. .553
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hW43f05"
/lab_host="JM107 or DHSa"
/clone_lib="MGS-Zmaysef (JM107 adapted methyl filtered)"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA

FEATURES

source

1. .553
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hW43f05"
/lab_host="JM107 or DHSa"
/clone_lib="MGS-Zmaysef (JM107 adapted methyl filtered)"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA

ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.5e-157; Mismatches 0; Indels 0; Gaps 0;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ACAACCTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 441
DB 442 ACACACTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 393
QY 442 TAGATATTATGATCATTTGAATCGTTTATCTCTTGAAGCGGTTTCAATTTTTTTTAC 501
DB 382 TAGATATTATGATCATTTGAATCGTTTATCTCTTGAAGCGGTTTCAATTTTTTTTAC 323
QY 502 AGACGCTTTTTTTAGAGGTGAGCATCATTAATGCGGATAGGTGTATACATCGGATTA 561
DB 322 AGACGCTTTTTTTAGAGGTGAGCATCATTAATGCGGATAGGTGTATACATCGGATTA 263
QY 562 CAACCTTAACCGTACACCACTTTTGAAGAGCTGTGAATGCGGATCTCTTCGCGTACCA 621
DB 262 CAACCTTAACCGTACACCACTTTTGAAGAGCTGTGAATGCGGATCTCTTCGCGTACCA 203
QY 622 GCACCTTTTACCAATCTTGTCTGCTGTGCAAAACCACTGTACGAATAGCATCTACT 681
DB 202 GCACCTTTTACCAATCTTGTCTGCTGTGCAAAACCACTGTACGAATAGCATCTACT 143
QY 682 GTTCT 686
DB 142 GTTCT 138

RESULT 3
LOCUS B2314790 648 bp DNA linear GSS 06-NOV-2002
DEFINITION h222a01.b1 MGS-Zmaysef (JM107 adapted methyl filtered) Zea mays
ACCESSION B2314790
VERSION B2314790.1 GI:24683733
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 648)
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,
Katzemburger, F., King, L., Miller, B., Muller, S., Nascento, L.,
Zutavern, T., McCombie, W.R., and Martensen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)

AUTHORS Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org

COMMENT Plate: h222 row: a column: 01
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 648.
Location/Qualifiers
1. .648
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="h222a01"

FEATURES

source

1. .648
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="h222a01"

was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(XbaI reads in M13mp19, BglI reads in pUC19). The same
ligation was transformed in either JM107 or DHSa. "

TITLE Genomic shotgun sequences from Zea mays (methyl-filtered)
JOURNAL Unpublished (2002)
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: h882 row: h column: 12
Seq primer: -21M13univFwd
Class: shotgun
High quality sequence stop: 776.

FEATURES
source
1. 776
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="h82h12"
/lab_host="JM107 or DH5a"
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(x/y reads in M13mp19. b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 776;
Best Local Similarity 100.0%; Pred. No. 2.5e-157;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ACACACTTCCACTGTTAGTGTGATGATCTGTTATCTTTCGAAACCATTAACAGAC 441
DB 197 ACACACTTCCACTGTTAGTGTGATGATCTGTTATCTTTCGAAACCATTAACAGAC 256
QY 442 TAGATATTTGATCATGTAATCGTTATTTCTGTGAAGCGGTTTCATTTTTTTTAC 501
DB 257 TAGATATTTGATCATGTAATCGTTATTTCTGTGAAGCGGTTTCATTTTTTTTAC 316
QY 502 AGACGCTTTTTTTAGAGGTGACATCCATTATGCGGCATATGATGTTACATCGCATTA 561
DB 317 AGACGCTTTTTTTAGAGGTGACATCCATTATGCGGCATATGATGTTACATCGCATTA 376
QY 562 CAACCTTAACCGTACACCATTTTGAAGATGCTGTGAATGCGGATCTTTCGGCTACCA 621
DB 377 CAACCTTAACCGTACACCATTTTGAAGATGCTGTGAATGCGGATCTTTCGGCTACCA 436
QY 622 GCACCTTTTACATTAATCTTGTGCTGTGCAAAACCATGTAAGATGATGATGCT 681
DB 437 GCACCTTTTACATTAATCTTGTGCTGTGCAAAACCATGTAAGATGATGATGCT 496
QY 682 GTTCT 686
DB 497 GTTCT 501

RESULT 8 824 bp mRNA linear EST 03-JUN-2003
CD438581/c
LOCUS E101N0515A02.b Endosperm_5 Zea mays cDNA, mRNA sequence.
DEFINITION CD438581
ACCESSION CD438581.1 GI:31354224
VERSION
KEYWORDS
SOURCE
ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 824)
AUTHORS Lai J., Dey N., Kim C.S., Bharti A.K., Rudd S., Mayer K.F.,
Laikins B., Becraft P. and Messing J.
TITLE Characterization of the maize endosperm transcriptome and its
JOURNAL comparison to the rice genome
COMMENT Genome Res. 14 (10), 1932-1937 (2004)
CONTACT: Lai, Jinheng
Dr. Joachim Messing's lab
Wakeman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@wakeman.rutgers.edu
Seq primer: T3

FEATURES
source
1. 824
Location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/cultiivar="W22"
/db_xref="taxon:4577"
/tissue_type="Endosperm_5"
/clone_lib="Endosperm_5"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI"

ORIGIN

Query Match 25.8%; Score 305; DB 6; Length 824;
Best Local Similarity 100.0%; Pred. No. 2.5e-157;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ACACACTTCCACTGTTAGTGTGATGATCTGTTATCTTTCGAAACCATTAACAGAC 441
DB 462 ACACACTTCCACTGTTAGTGTGATGATCTGTTATCTTTCGAAACCATTAACAGAC 403
QY 442 TAGATATTTGATCATGTAATCGTTATTTCTGTGAAGCGGTTTCATTTTTTTTAC 501
DB 402 TAGATATTTGATCATGTAATCGTTATTTCTGTGAAGCGGTTTCATTTTTTTTAC 343
QY 502 AGACGCTTTTTTTAGAGGTGACATCCATTATGCGGCATATGATGTTACATCGCATTA 561
DB 342 AGACGCTTTTTTTAGAGGTGACATCCATTATGCGGCATATGATGTTACATCGCATTA 283
QY 562 CAACCTTAACCGTACACCATTTTGAAGATGCTGTGAATGCGGATCTTTCGGCTACCA 621
DB 282 CAACCTTAACCGTACACCATTTTGAAGATGCTGTGAATGCGGATCTTTCGGCTACCA 223
QY 622 GCACCTTTTACATTAATCTTGTGCTGTGCAAAACCATGTAAGATGATGATGCT 681
DB 222 GCACCTTTTACATTAATCTTGTGCTGTGCAAAACCATGTAAGATGATGATGCT 163
QY 682 GTTCT 686
DB 162 GTTCT 158

RESULT 9 920 bp DNA linear GSS 16-MAY-2003
CC357906
LOCUS PNHU36TD ZW.0.6.1.0 KB Zea mays genomic clone ZMWBRta454F23,
DEFINITION CC357906
ACCESSION CC357906
VERSION CC357906.1 GI:30827306
KEYWORDS
SOURCE
ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 920)
AUTHORS Whitelaw C.A., Quackenbush J., Van Aken S., Uterback T.,
Reinick A., Fraser C.M., Yuan Y., San Miguel P., Ma J. and
Benutzer J.
TITLE Maize Genomics Consortium

Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcmombe@cshl.org
Plate: 1678 row: f column: 08
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 622.
Location/Qualifiers

FEATURES
source

1. 600
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="1678F08"
/lab_host="DH5a"
/clone_11b="WGS-ZmaysF (DH5a methyl filtered)"
/note="Organ: Immature ears; Site 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (x/y
reads in M13mp19, .b/g reads in pUC19). The same ligation
was transformed into DH5a."

ORIGIN

Query Match 23.2%; Score 275; DB 8; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.2e-140; Indels 0; Gaps 0;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 CCTGTATCTCTTCGACCATTAACAGACTAGTATTATTTGATCATTTGATTCGTTATT 471
DB 1 CCTGTATCTCTTCGACCATTAACAGACTAGTATTATTTGATCATTTGATTCGTTATT 60
QY 472 TCTCTTGAAGCGGTTTCATTTTATTTTACAGAGCTCTTTTATAGAGTTCGACATCCA 531
DB 61 TCTCTTGAAGCGGTTTCATTTTATTTTACAGAGCTCTTTTATAGAGTTCGACATCCA 120
QY 532 TTATGCGGCATAGGTGTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCAATG 591
DB 121 TTATGCGGCATAGGTGTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCAATG 180
QY 592 GCTGTAATGCGGATCTCTCCGCTACACGACCTTTTACCATTAACCTTCTGCTGTTGC 651
DB 181 GCTGTAATGCGGATCTCTCCGCTACACGACCTTTTACCATTAACCTTCTGCTGTTGC 240
QY 652 AAACCCACTGTACGATATGATCTACTGCTGTTCT 686
DB 241 AAACCCACTGTACGATATGATCTACTGCTGTTCT 275

RESULT 12
CC025412/c 313 bp DNA linear GSS 01-APR-2003
LOCUS 3591.1.46.1 A05.1E1.X.1 3591 - RescueMu Grid P Zea mays genomic.
DEFINITION genomic survey sequence.

ACCESSION CC025412.1 GI:29440269
VERSION CC025412.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 313)

REFERENCE
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University

855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 3591.1.46.1 row: 21
Class: transposon-tagged.
Location/Qualifiers

FEATURES
source

1. 313
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_11b="3591 - RescueMu Grid P"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.jastate.edu' and follow the links for
'RescueMu.' Grid P was grown at Molokai in 2002. DNA was
extracted from leaf strips, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."

ORIGIN

Query Match 23.1%; Score 273; DB 8; Length 313;
Best Local Similarity 100.0%; Pred. No. 1.5e-139; Indels 0; Gaps 0;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 409 GATCCGTATCTCTTCGACCATTAACAGACTAGTATTATTTGATCATTTGATTCGTTATT 468
DB 273 GATCCGTATCTCTTCGACCATTAACAGACTAGTATTATTTGATCATTTGATTCGTTATT 214
QY 469 ATTTCTCTTGAAGCGGTTTCATTTTATTTTACAGAGCTCTTTTATAGAGTTCGACAT 528
DB 213 ATTTCTCTTGAAGCGGTTTCATTTTATTTTACAGAGCTCTTTTATAGAGTTCGACAT 154
QY 529 CCATTATGCGGATAGGTGTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCA 588
DB 153 CCATTATGCGGATAGGTGTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCA 94
QY 589 ATGCGTGTATAGCGGATCTCTCCGCTACACGACCTTTTACCATTAACCTTCTGCTGCT 648
DB 93 ATGCGTGTATAGCGGATCTCTCCGCTACACGACCTTTTACCATTAACCTTCTGCTGCT 34
QY 649 TGCAAAACCACTGTACGATATGATCTACTGCT 681
DB 33 TGCAAAACCACTGTACGATATGATCTACTGCT 1

RESULT 13
B2359336 646 bp DNA linear GSS 18-NOV-2002
LOCUS 1d72d02.b1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
DEFINITION genomic clone 1d72d02 5', genomic survey sequence.

ACCESSION B2359336
VERSION B2359336.1 GI:25059030
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 646)

REFERENCE
AUTHORS Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,

TITLE
JOURNAL
COMMENT

Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: 1d72 row: d column: 02
Seq primer: -21M13univFwd
Class: Shotgun
High quality sequence stop: 646.
Location/Qualifiers

FEATURES
source

1. 646
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="1d72d02"
/lab_host="JM107 or DH5a"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(X/Y reads in M13mp19, B/G reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

ORIGIN

Query Match 21.6%; Score 256; DB 8; Length 646;
Best Local Similarity 100.0%; Pred. No. 4.2e-130; Indels 0; Gaps 0;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ACACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGAC 441
DB 361 ACACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGAC 420
QY 442 TAGATATTATGATCATGATCGTTATTTCTCTTGAAGCGGTTTCATTTTTTTTAC 501
DB 421 TAGATATTATGATCATGATCGTTATTTCTCTTGAAGCGGTTTCATTTTTTTTAC 480
QY 502 AGAGCTCTTTTATGAGAGTGCACATCATTTATGCGGCATAGGTTCATCGCGTAA 561
DB 481 AGAGCTCTTTTATGAGAGTGCACATCATTTATGCGGCATAGGTTCATCGCGTAA 540
QY 562 CAACCTTAACCGTACACCACTTTAGCAATGGCTCGTAATGGCGCATCTCTTCGCTACCA 621
DB 541 CAACCTTAACCGTACACCACTTTAGCAATGGCTCGTAATGGCGCATCTCTTCGCTACCA 600
QY 622 GCACCTTTTACCATTA 637
DB 601 GCACCTTTTACCATTA 616

RESULT 14
B2372017 650 bp DNA linear GSS 26-NOV-2002
LOCUS B2372017
DEFINITION 1662b08.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone
B2372017
ACCESSION B2372017.1 GI:25456191
VERSION B2372017.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 650)
AUTHORS Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,

TITLE
JOURNAL
COMMENT

Katzenburger, F., King, L., Miller, B., Muller, S., Nascimben, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: 1662 row: b column: 08
Seq primer: -21M13univFwd
Class: Shotgun
High quality sequence stop: 651.
Location/Qualifiers

FEATURES
source

1. 650
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="1662b08"
/lab_host="DH5a"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (X/Y
reads in M13mp19, B/G reads in pUC19). The same ligation
was transformed into DH5a."

ORIGIN

Query Match 21.4%; Score 253; DB 8; Length 650;
Best Local Similarity 99.7%; Pred. No. 1.9e-128; Indels 1; Gaps 0;
Matches 303; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 383 CACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGACT 442
DB 1 CACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGACT 60
QY 443 AGATATTATGATCATGATCGTTATTTCTCTTGAAGCGGTTTCATTTTTTTTAC 502
DB 61 AGATATTATGATCATGATCGTTATTTCTCTTGAAGCGGTTTCATTTTTTTTAC 480
QY 503 GACGCTCTTTTATGAGAGTGCACATCATTTATGCGGCATAGGTTCATCGCGTAA 562
DB 121 GACGCTCTTTTATGAGAGTGCACATCATTTATGCGGCATAGGTTCATCGCGTAA 540
QY 563 AACTTAACCGTACACCACTTTAGCAATGGCTCGTAATGGCGCATCTCTTCGCTACCA 622
DB 181 AACTTAACCGTACACCACTTTAGCAATGGCTCGTAATGGCGCATCTCTTCGCTACCA 600
QY 623 CACCTTTTACCATTAACCTTTGCTGTTGCAAGCCGCTGATGATGATGATGATGAT 682
DB 241 CACCTTTTACCATTAACCTTTGCTGTTGCAAGCCGCTGATGATGATGATGATGAT 660
QY 683 TTCT 686
DB 301 TTCT 304.

RESULT 15
B2363755 428 bp DNA linear GSS 18-NOV-2002
LOCUS B2363755
DEFINITION 1d90d01.g1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
B2363755 genomic clone 1d90d01 5', genomic survey sequence.
ACCESSION B2363755.1 GI:25068831
VERSION B2363755.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 428)

TITLE
JOURNAL
COMMENT

Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
Zutavern, T., McCombie, W.R. and Marienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)

Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: id90 row: d column: 01
Seq primer: -21M3UnivRev
Class: Shotgun
High quality sequence stop: 428.
Location/Qualifiers

FEATURES
source

1..428
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="id90d01"
/lab_host="JMI07 or DHSa"
/clone_lib="MGS-ZmaysF (JMI07 adapted methyl filtered)"
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(x/y reads in M13mp19, b/g reads in pUC19). The same
ligation was transformed in either JMI07 or DHSa."

ORIGIN

Query Match 20.4%; Score 241; DB 8; Length 428;
Best Local Similarity 99.7%; Pred. No. 9e-122;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 395 CTAGTGTGAGTGATCTGTATCTTCTCGAACCAATACAGACTAGTATTGGA 454
DB 1 CTAGTGTGAGTGATCTGTATCTTCTCGAACCAATACAGACTAGTATTGGA 60
QY 455 TCATTGAATCGTTATTTCTCTGAAAGCGTTTCATTTTTTTTACAGAGCTTTT 514
DB 61 TCATTGAATCGTTATTTCTCTGAAAGCGTTTCATTTTTTTTACAGAGCTTTT 120
QY 515 TAGGAGTCGATCATTTATGCGGATAGGTATACATCGGATATACACTTAACCGTA 574
DB 121 TAGGAGTCGATCATTTATGCGGATAGGTATACATCGGATATACACTTAACCGTA 180
QY 575 CACCACTTTAGCAATGCTGTAAATGCGGATCTTTCCGTTACAGACCTTTTACCA 634
DB 181 CACCACTTTAGCAATGCTGTAAATGCGGATCTTTCCGTTACAGACCTTTTACCA 240
QY 635 TAACTTGTGCTGTGCAAAACCACTGTACGAATAGCATCTACTGTTCT 686
DB 241 TAACTTGTGCTGTGCAAAACCACTGTACGAATAGCATCTACTGTTCT 292

Search completed: February 10, 2005, 17:01:49
Job time : 8001.05 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 117.196 Seconds
(without alignments)
7855.666 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19
Sequence: 1 TGTAGCGGCCACGCGTGG 19

Scoring table: OLIGO-NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	AX342370	Sequence
2	19	100.0	498	AX342368	Sequence
3	18	94.7	3178	AX492266	Coryneb
4	16	84.2	189	AY271325	AY271325
5	16	84.2	189	AY271340	AY271340
6	16	84.2	204	AY271897	AY271897
7	16	84.2	3180	AY492278	AY492278
8	16	84.2	3330	AY492261	AY492261
9	16	84.2	3333	AY492259	AY492259
10	16	84.2	3357	AY492234	AY492234
11	16	84.2	3357	AY492285	AY492285
12	16	84.2	3702	AY147167	AY147167
13	16	84.2	3868	AY147173	AY147173
14	16	84.2	3906	AY262736	AY262736
15	16	84.2	3929	AY147165	AY147165
16	16	84.2	4405	SC0938104	SC0938104
17	16	84.2	299050	AP005027	AP005027
18	16	84.2	301675	AC092285	AC092285
19	16	84.2	316296	2	AC092285

C	20	15	78.9	1183	10	RATREP3B	D16443 Rattus norv
C	21	15	78.9	1211	10	RNEPB3	X80133 R.norvegicu
C	22	15	78.9	1253	10	RATREP3R	D29969 Rat mRNA fo
C	23	15	78.9	1275	10	AF302686	AF302686 Rattus no
C	24	15	78.9	1310	10	RATREP3	D14869 Rattus norv
C	25	15	78.9	1345	10	RNEPB3ALPH	X83855 R.norvegicu
C	26	15	78.9	1405	6	AR038125	AR038125 Sequence
C	27	15	78.9	1405	6	I28939	I28939 Sequence 3
C	28	15	78.9	1616	8	AK111750	AK111750 Oryza sat
C	29	15	78.9	1974	4	AF266477	AF266477 Canis fam
C	30	15	78.9	2002	10	MUSPGERB	D13321 Mus musculu
C	31	15	78.9	2006	6	E06015	E06015 DNA encodin
C	32	15	78.9	2090	10	BC058742	BC058742 Mus muscu
C	33	15	78.9	2091	10	MUSPGER	D10204 Mus musculu
C	34	15	78.9	2107	6	AR038124	AR038124 Sequence
C	35	15	78.9	2107	6	E06014	E06014 DNA encodin
C	36	15	78.9	2107	6	I28938	I28938 Sequence 1
C	37	15	78.9	2556	14	HSVTKY	K02122 Matmoselet he
C	38	15	78.9	2811	5	GGU47273	U47273 Gallus gall
C	39	15	78.9	3245	5	CHKTRANGLU	L02270 Chicken tra
C	40	15	78.9	3570	10	MUSPGERB	D17406 Mus musculu
C	41	15	78.9	6219	1	AF340167	AF340167 Streptomy
C	42	15	78.9	7000	10	AF458960	AF458960 Mus muscu
C	43	15	78.9	15674	9	AC118277	AC118277 Homo sapi
C	44	15	78.9	51855	1	ML1557546	AJ557546 Melittang
C	45	15	78.9	52684	8	AF111709	AF111709 Oryza sat

ALIGNMENTS

RESULT 1	AX342370	19 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342370	Sequence 9 from Patent EP1167531.			
DEFINITION	AX342370				
ACCESSION	AX342370.1	GI:18151813			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE	1	Behr, C.F., Hironaka, C., Heck, G.R. and You, J.			
AUTHORS		Corn transformant pv-zmgc32 (nke03) and compositions and methods			
TITLE		for detection thereof			
JOURNAL		Patent: EP 1167531-A 9 02-JAN-2002;			
FEATURES		Monsanto Technology LLC (US)			
SOURCE		Location/Qualifiers			
		1..19			
		/organism="synthetic construct"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:32630"			
		/note="Zea maize genomic and vector DNA"			
ORIGIN					
Query Match	100.0%	Score 19;	DB 6;	Length 19;	
Best Local Similarity	100.0%	Pred. No. 24;			
Matches	19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Cy	1	TGTAGCGGCCACGCGTGG	19		
Db	1	TGTAGCGGCCACGCGTGG	19		
RESULT 2	AX342368	498 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342368	Sequence 7 from Patent EP1167531.			
DEFINITION	AX342368				
ACCESSION	AX342368.1	GI:18151811			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
		synthetic construct			
		synthetic construct			

other sequences; artificial sequences.

REFERENCE

1 Behr, C.F., Hironaka, C., Heck, G.R. and You, J.
Corn transformatant PV-zmgpt32 (nk603) and compositions and methods
for detection thereof
Patent: EP 1167531-A 7 02-JAN-2002;
Monsanto Technology LLC (US)

JOURNAL

FEATURES

source

1. 498
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="1-304 Zea maize genomic DNA 305-349 construct
vector DNA 350-498 rice actin 1 promoter DNA"

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 498;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 TGTAGCGCGCCACGCGTG 19
|||||
Db 296 TGTAGCGCGCCACGCGTG 314

RESULT 3

AY492266 3178 bp DNA linear BCT 14-SEP-2004

LOCUS

Corynebacterium afermentans subsp. lipophilum strain CIP 103500

DEFINITION

AY492266

ACCESSION

AY492266

VERSION

AY492266.1 GI:46360989

KEYWORDS

ORGANISM

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

13385, France

Location/Qualifiers

1. 3178

/organism="Corynebacterium afermentans subsp. lipophilum"

/mol_type="genomic DNA"

/strain="CIP 103500"

/sub_species="lipophilum"

/db_xref="taxon:144184"

/clone="carpob37"

/note="type strain of Corynebacterium afermentans subsp. lipophilum"

<1. .>3178

/gene="rpoB"

<1. .>3178

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RpoB"

/protein_id="AAS89228.1"

/db_xref="GI:46360990"

/translation="TAAVSDOTMNAADPFGAPERYSPAKINPITVPGLLDVOLSPFAV
LVGQEMREBERGAGNADARTSGLEIDLEISIPIDYSGMGLTISPREEDVKYTI
DECKDKINYSAPLYVTAFINDDTOKSTVITGPPMTDGTITVNTKTERVYVS
QIVRSFGVYDEITDKSTERPLSHVKVIPSAGAWLEFVDVDRDVGAVIDKKRQPV
VLKALGWTTEQITERGFEISIMSTLENDVSNIDEALTEIKRQREGQPTDLAQ
SLIENSPFAKRYDLAVGRYKTRKLGIGDHDGLMTLTEDTATLTLEYIVRLHAG

ORIGIN

Query Match 94.7%; Score 18; DB 1; Length 3178;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 TGTAGCGCGCCACGCGTG 18
|||||
Db 892 TGTAGCGCGCCACGCGTG 875

RESULT 4

AY271325

189 bp DNA linear BCT 07-MAY-2003

LOCUS

Mycobacterium fortuitum type I RNA polymerase beta subunit (rpoB)

gene, partial cds.

AY271325

AY271325.1 GI:30145567

KEYWORDS

ORGANISM

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

12734283

2 (bases 1 to 189)

Bang, H.-E., Park, H.-J., Bai, G.-H., Cho, S.-N. and Lee, H.

Submitted (03-APR-2003) Biomedical Laboratory Science, College of

Health Sciences, Maeji-Ri, Heungup-Myun, Wonju, Kangwon-do 220-710,

Korea

Location/Qualifiers

1. 189

/organism="Mycobacterium fortuitum"

/mol_type="genomic DNA"

/db_xref="taxon:1766"

/note="type: I"

<1. .>189

/gene="rpoB"

<1. .>189

/gene="rpoB"

/codon_start=2

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="AAP15214.1"

/db_xref="GI:30145568"

/translation="LAAVGRYKYNKKGLGAAAGPITSSITLDEVDVATIEYVRLHAG
QTTWTVGVEVEVD"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. No. 4,4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 TGTAGCGCGCCACGCG 16

Db 23 TGTAGCGGCCACGCG 8

RESULT 5
LOCUS AY271340/c 189 bp DNA linear BCT 07-MAY-2003
DEFINITION Mycobacterium fortuitum type II RNA polymerase beta subunit (rpoB)
ACCESSION AY271340
VERSION AY271340.1 GI:30145597
KEYWORDS
SOURCE
ORGANISM Mycobacterium fortuitum
Mycobacterium fortuitum
Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

REFERENCE
AUTHORS Lee, H., Bang, H. E., Bai, G. H. and Cho, S. N.
TITLE Novel Polymorphic Region of the rpoB Gene Containing Mycobacterium Species-Specific Sequences and Its Use in Identification of Mycobacteria
J. Clin. Microbiol. 41 (5), 2213-2218 (2003)
12734283

JOURNAL
PUBMED 2 (bases 1 to 189)
REFERENCE Bang, H. E., Park, H. J., Bai, G. H., Cho, S. N. and Lee, H.
AUTHORS Direct Submission
TITLE Submitted (03-APR-2003) Biomedical Laboratory Science, College of Health Sciences, Maeji-Ri, Heungup-Myun, Wonju, Kangwon-Do 220-710, Korea

FEATURES
source
1..189
/organism="Mycobacterium fortuitum"
/mol_type="genomic DNA"
/db_xref="taxon:1766"
/note="type: II"
<1..>189
/gene="rpoB"
<1..>189
/gene="rpoB"
/codon_start=2
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="AAPI5229.1"
/db_xref="GI:30145598"
/translation="LAVRGYKVKKLGKLNAGDPITSTLTEDVATIEYLRLHMG QTMVTPGVEVPVETDD"

ORIGIN
Query Match 84.2%; Score 16; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCG 16
|||||
23 TGTAGCGGCCACGCG 8

Db 23 TGTAGCGGCCACGCG 8

RESULT 6
LOCUS AY271897/c 204 bp DNA linear BCT 07-MAY-2003
DEFINITION Mycobacterium kansasii type III RNA polymerase beta subunit (rpoB)
ACCESSION AY271897
VERSION AY271897
KEYWORDS
SOURCE
ORGANISM Mycobacterium kansasii
Mycobacterium kansasii
Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

REFERENCE
AUTHORS Lee, H., Bang, H. E., Bai, G. H. and Cho, S. N.
TITLE Novel Polymorphic Region of the rpoB Gene Containing Mycobacterium Species-Specific Sequences and Its Use in Identification of

JOURNAL
PUBMED 12734283
REFERENCE Bang, H. E., Park, H. J., Bai, G. H., Cho, S. N. and Lee, H.
AUTHORS Direct Submission
TITLE Submitted (10-APR-2003) Biomedical Laboratory Science, College of Health Sciences, Maeji-Ri, Heungup-Myun, Wonju, Kangwon-Do 220-710, Korea

FEATURES
source
1..204
/organism="Mycobacterium kansasii"
/mol_type="genomic DNA"
/db_xref="taxon:1768"
/note="type: III"
<1..>204
/gene="rpoB"
<1..>204
/gene="rpoB"
/codon_start=2
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="AAPI5235.1"
/db_xref="GI:30145610"
/translation="LAVRGYKVKKLGKLNAGDPITSTLTEDVATIEYLRLHHA SDQGPANWTPGVEVPVETDD"

ORIGIN
Query Match 84.2%; Score 16; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCG 16
|||||
23 TGTAGCGGCCACGCG 8

Db 23 TGTAGCGGCCACGCG 8

RESULT 7
LOCUS AY492278/c 3180 bp DNA linear BCT 14-SEP-2004
DEFINITION Corynebacterium riegellii strain CIP 105310 rpoB (rpoB) gene, partial cds.
ACCESSION AY492278
VERSION AY492278.1 GI:46361013
KEYWORDS
SOURCE
ORGANISM Corynebacterium riegellii
Corynebacterium riegellii
Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS Khamis, A., Raoult, D. and La Scola, B.
TITLE rpoB Gene Sequencing for Identification of Corynebacterium Species
J. Clin. Microbiol. 42 (9), 3925-3931 (2004)
15364970

JOURNAL
PUBMED 2 (bases 1 to 3180)
REFERENCE La Scola, B., Khamis, A. and Raoult, D.
AUTHORS Direct Submission
TITLE Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48, Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille 13385, France

FEATURES
source
1..3180
/organism="Corynebacterium riegellii"
/mol_type="genomic DNA"
/strain="CIP 105310"
/db_xref="taxon:156976"
/clone="crpob49"
/note="type strain of Corynebacterium riegellii"
<1..>3180
/gene="rpoB"
<1..>3180
/gene="rpoB"
/codon_start=1

/tranel_table=11
/product="RpoB"
/protein_id="AAS89240.1"
/db_xref="GI:46361014"
/translation="LANSDDTMSMAEIPGAPERYSPAKINEPIVPGLLDVOLLESFAM
LVGTSEMERREOQLRGFTARVTSGLDELIEISPIQDYSGNMSTLSSEPRDEKXTTI
BEKKDKDINYSAPLYTAIEFIINDTOEIKSQTVIIGDFPLMTEKGTPIVNGTERVVS
QVLRSPGVYDEFTIDKSTERPLHSVKYIPSKGAMLEFDVDRDVGVRIDKRRQPTV
VLKALGWTTEQITERFSEIMMATTLESQVANTDALLEIYKORQGEQPTDLAQ
SILENSFFKAKRQDLARVGRYKVRKLGSDHGLMTLEDEIATTELYVRLAHAB
SEMTSPSGEIIPISTDDIDHGRRLRTVTEGILIONOVYGLSPMERVREMTTODAE
SITPISLINRVSAIREFGTSLSQFMDONNSLSGLTHKRRLSLAGRCSTREPA
GIEVRDHPSHYGMCPLETPPEGNIIGLALASVAVNPGFIETPYQKNDKLTID
QVYLTADBEDRYAIAQAATPMDKGNLTGERIEVRLKDGDIYVVGKGYDYLISPR
QMSVATAMIPLIEDDANRALMGANMOQKAVPILRAEAAYATGMEORAAVYDADTV
ISPRAGVETVGTQYITVMDDEGRDVTMLRTPERTNOCTYNOTPIVSGDRVEAGQ
VIADPGCTKGEWALGRNLLVAPNPMGHTYEDLIINQRYVEEDILTSVHIEBEDD
ARDTKGAEIETREIPNVEDVLKDLBERGIRIGADVRDGDILVGRVTPKGETELP
EERLLRAIFEEKAREVDRDTSIKVPHSQGVIAVRRFARDEDDDLARQVEMIRVYA
QKRIKIDGDQAGHGNKGVKILPOEDMFPNADGTPVDIILNTHGVPFRMNTIGVLA
EVLGLAKAKAGTVPDPKNAKLETLPEHLVYVPAVNSLTATPFDGATVDEIAGLL
ANSKPNRDGVMDENGKTMLPDGRSGEPYVPIISGVYMWLKLHLVDEKIHARSNG
PYSMITQOPLGKRAQFGQGRFGEHNEWAMQ"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3180;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCGG 16
Db 892 TGTAGCGGCCCGCGG 877

RESULT 8 3330 bp DNA linear BCT 14-SEP-2004
AY492261/c
LOCUS
DEFINITION Corynebacterium mucifaciens strain CIP 105129 RpoB (rpoB) gene,
partial cds.
ACCESSION AY492261
VERSION AY492261.1 GI:46360979
KEYWORDS
SOURCE
ORGANISM Corynebacterium mucifaciens
Corynebacterium mucifaciens
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS 1 (bases 1 to 3330)
TITLE Khamis, A., Raoult, D. and La Scola, B.
JOURNAL rpoB Gene Sequencing for Identification of Corynebacterium Species
PUBMED 15364970
AUTHORS 2 (bases 1 to 3330)
TITLE J. Clin. Microbiol. 42 (9), 3925-3931 (2004)
REFERENCE
AUTHORS La Scola, B., Khamis, A. and Raoult, D.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille
13385, France

FEATURES
source

1..3330 Location/Qualifiers
/organism="Corynebacterium mucifaciens"
/mol_type="genomic DNA"
/strain="CIP 105129"
/db_xref="taxon:57171"
/clone="cmrpb32"
/note="type strain of Corynebacterium mucifaciens"
<1..3330
/gene="rpoB"
<1..3330
/codon_start=3
/tranel_table=11
/product="RpoB"
/protein_id="AAS89223.1"

/db_xref="GI:46360980"
/translation="SRQTNMADIPGAPERYSPAKINEPIVPGLLDVOLLESFAMLVG
TPMERREOQLRGFTARVTSGLDELIEISPIEDYSGNMSLTSSEPRDEKXTTIDEC
KDKDINYSAPLYTAIEFIINDTOEIKSQTVIIGDFPLMTEKGTPIVNGTERVVSQV
LRSPGVYDEFTIDKSTERPLHSVKYIPSKGAMLEFDVDRDVGVRIDKRRQPTVIL
KALGWTTEQITERFSEIMMATTLESQVANTDALLEIYKORQGEQPTDLAQSL
ENSFPAKRDYDLARVGRYKVRKLGSDHGLMTLEDEIATTELYVRLAHABSEM
TSPEGELIPIINTDDIDHGRRLRTVTEGILIONOVYGLSPMERVREMTTODAESIT
PTSLINRVSAIREFGTSLSQFMDONNSLSGLTHKRRLSLAGRCSTREPA
VVRVSHYGMCPLETPPEGNIIGLALASVAVNPGFIETPYQKNDKLTID
YLTADBEDRYAIAQAATPMDKGNLTGERIEVRLKDGDIYVVGKGYDYLISPRQV
SVATAMIPLIEDDANRALMGANMOQKAVPILRAEAAYATGMEORAAVYDADTVISK
KAGVETVGTQYITVMDDEGRDVTMLRTPERTNOCTYNOTPIVSGDRVEAGQVIA
DGEQTKGEWALGRNLLVAPNPMGHTYEDLIINQRYVEEDILTSVHIEBEDD
TKGAEIETREIPNVEDVLKDLBERGIRIGADVRDGDILVGRVTPKGETELP
LRAIFEEKAREVDRDTSIKVPHSQGVIAVRRFARDEDDDLARQVEMIRVYA
KIDGQKMAGRHGNKGVKILPOEDMFPNADGTPVDIILNTHGVPFRMNTIGVLEI
LGLAKAGTVPDPKNAKLETLPEHLVYVPAVNSLTATPFDGATVDEIAGLLANS
KPNRDGVMDENGKTMFPDGRSGEPYVPIISGVYMWLKLHLVDEKIHARSNG
MYQOPLGKRAQFGQGRFGEHNEWAMQAYATYQELTIKSDVGRKXYEAIIVK
GNIIPGQIPESFKVLKEIQ"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3330;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCGG 16
Db 885 TGTAGCGGCCCGCGG 870

RESULT 9 3333 bp DNA linear BCT 14-SEP-2004
AY492259/c
LOCUS
DEFINITION Corynebacterium imitans strain CIP 105130 RpoB (rpoB) gene, partial
cds.
ACCESSION AY492259
VERSION AY492259.1 GI:46360975
KEYWORDS
SOURCE
ORGANISM Corynebacterium imitans
Corynebacterium imitans
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS 1 (bases 1 to 3333)
TITLE Khamis, A., Raoult, D. and La Scola, B.
JOURNAL rpoB Gene Sequencing for Identification of Corynebacterium Species
PUBMED 15364970
AUTHORS 2 (bases 1 to 3333)
TITLE J. Clin. Microbiol. 42 (9), 3925-3931 (2004)
REFERENCE
AUTHORS La Scola, B., Khamis, A. and Raoult, D.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille
13385, France

FEATURES
source

1..3333 Location/Qualifiers
/organism="Corynebacterium imitans"
/mol_type="genomic DNA"
/strain="CIP 105130"
/db_xref="taxon:156978"
/clone="clrpb30"
/note="type strain of Corynebacterium imitans"
<1..3333
/gene="rpoB"
<1..3333
/codon_start=3
/tranel_table=11
/product="RpoB"
/protein_id="AAS89231.1"
/db_xref="GI:46360976"
/translation="SRQTNMADIPGAPERYSPAKINEPIVPGLLDVOLLESFAMLVG

TEEMREBOELGDEARVKSGLIEDILDEISPIODYSGNMSLTISEPREDAKYTBEE
KOKDINYSAPLYVTAEPINNDQELKSQTVFICGDFPLMTDKGTFVINGTERVVSOLV
RSGVYFDEITDKSTERPLHAVKVIIPSRGAMLEFDVKDRDVTGVRIDRRKORPEO
KALGFTTBOITERFSELMSTLNDGVAMTEDEALLETIRKORPEOETPRDIAOSIL
ENSPFKARVDIARVGRYKVRNKLGLGSDHEDMLTEEDIALTEYLVRILHAGETEM
TSPSGGIIITINDDDIHFGNRLRTVYGEILQNOVRGLSRMEVNERMTDOAESIT
PSTLINRVPSAIREFFGTSQSDPMDNLSGLTHRRLSALCPGSLSRRAIE
VADVHPSHYGRMCPIETEPGPNIGLIGALSAYARVNFEGITETPRKVDNGLTDID
YLTADEEDRYAIAQAATPMTKDELGERIEVLKDGIDGVGPOGVYLDISPRQVY
SVATAMIPELHEDDNRALMAGAMHQAVALREAPYATGMEORAAADAGDTVISP
LGVATVTVGDTYTVIGDGGTRDPTQLRPFHRTNOGTCYNOPTVISOGEVAGOVIA
DEPGTODGMSLGRNLLVAFMPEBGNRYEDAILNORVYBEDILTSVHIEHEIDARD
TKLGAETIREIPNVSEDLKDLDEKGIIRIGADVDDGDI LVKVTPKGTELTPEER
LRLAIFGEKAREVDRDTSLVKPHGETSKVIAVRFSREDDDLSPGVNEMIRVVAQKR
KIDODGKMGARHNGNKGVGKILPTEBDMPEMEDGTPVDIILNTHGVPRRNMIGQVLEH
LGMILAKAGMTVNDPDPANAKELIETLEPEKLYDVPESLTATPVFGASNEBIAGLLANS
KPRRDGVDVWNGDGKARLDGSGEPKPYKPVSGVYMWMLKHLHLDKIHARSSTGPS
MITQDPLGKAQGGORFGEMERWMAQAYGAATTLQDELLTIKSDDVVGRKYVEALVK
GDNIPDPGIPESFKVLKELQS"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3333;

Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCCCACGCG 16

Db 885 TGTAGCGGCCCCACGCG 870

RESULT 10
AY492234/c 3357 bp DNA linear BCT 14-SEP-2004
LOCUS
DEFINITION
Corynebacterium auris strain CIP 104632 RpoB (rpoB) gene, partial

ACCESSION
AY492234
VERSION
KEYWORDS
SOURCE
ORGANISM

Corynebacterium auris
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
1 (bases 1 to 3357)

AUTHORS
Khamis, A., Raoult, D. and La Scola, B.
TITLE
rpoB Gene Sequencing for Identification of Corynebacterium Species

JOURNAL
J Clin. Microbiol. 42 (9), 3925-3931 (2004)

PUBMED
15364970

REFERENCE
2 (bases 1 to 3357)

AUTHORS
La Scola, B., Khamis, A. and Raoult, D.
TITLE
Direct Submision
Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille
13385, France

FEATURES
source

1. .3357
Location/Qualifiers
/organism="Corynebacterium auris"
/mol_type="genomic DNA"
/strain="CIP 104632"
/db_xref="taxon:44750"
/clone="carp05"
/note="type strain of Corynebacterium auris"
<1. .3357
/gene="rpoB"
<1. .3357
/gene="rpoB"
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="rpoB"
/protein_id="AAS89196.1"
/db_xref="GI:46360926"
/translation="SRQMSMATTPGAPORVSPAKISRPITVPGLLDVORSPFALV
GSPWREDAARAGBGRVTSIGLEIDLELSPIDYSGNMSLSISPRREDYKTTIDEC
KDMTYAALFVTAEPINNDQELKSQTVFICGDFPLMTDKGTFVINGTERVVSOLV
KOKDINYSAPLYVTAEPINNDQELKSQTVFICGDFPLMTDKGTFVINGTERVVSOLV

RSGVYFDEITDKSTERPLHAVKVIIPSRGAMLEFDVKDRDVTGVRIDRRKORPEO
KALGFTTBOITERFSELMSTLNDGVAMTEDEALLETIRKORPEOETPRDIAOSIL
ENAPFRARVDIARVGRYKVRNKLGLGSDHEDMLTEEDIALTEYLVRILHAGETEM
TSPNGAVVINDDDIHFGNRLRTVYGEILQNOVRGLSRMEVNERMTDOAESIT
PSTLINRVPSAIREFFGTSQSDPMDNLSGLTHRRLSALCPGSLSRRAIE
VADVHPSHYGRMCPIETEPGPNIGLIGALSAYARVNFEGITETPRKVDNGLTDID
YLTADEEDRYAIAQAATPMTKDELGERIEVLKDGIDGVGPOGVYLDISPRQVY
SVATAMIPELHEDDNRALMAGAMHQAVALREAPYATGMEORAAADAGDTVISP
LGVATVTVGDTYTVIGDGGTRDPTQLRPFHRTNOGTCYNOPTVISOGEVAGOVIA
DEPGTODGMSLGRNLLVAFMPEBGNRYEDAILNORVYBEDILTSVHIEHEIDARD
TKLGAETIREIPNVSEDLKDLDEKGIIRIGADVDDGDI LVKVTPKGTELTPEER
LRLAIFGEKAREVDRDTSLVKPHGETSKVIAVRFSREDDDLSPGVNEMIRVVAQKR
KIDODGKMGARHNGNKGVGKILPTEBDMPEMEDGTPVDIILNTHGVPRRNMIGQVLEH
LGMILAKAGMTVNDPDPANAKELIETLEPEKLYDVPESLTATPVFGAETNEBIAGLLANS
KPRRDGVDVWNGDGKARLDGSGEPKPYKPVSGVYMWMLKHLHLDKIHARSSTGPS
MITQDPLGKAQGGORFGEMERWMAQAYGAATTLQDELLTIKSDDVVGRKYVEALVK
GDNIPDPGIPESFKVLKELQSCLNVEVL"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3357;

Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCCCACGCG 16

Db 885 TGTAGCGGCCCCACGCG 870

RESULT 11
AY492285 3357 bp DNA linear BCT 14-SEP-2004
LOCUS
DEFINITION
Rhodococcus equi strain CIP 81.17 RpoB (rpoB) gene, partial cds.

ACCESSION
AY492285
VERSION
KEYWORDS
SOURCE
ORGANISM

Rhodococcus equi
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Nocardiaceae; Rhodococcus.

REFERENCE
1 (bases 1 to 3357)

AUTHORS
Khamis, A., Raoult, D. and La Scola, B.
TITLE
rpoB Gene Sequencing for Identification of Corynebacterium Species

JOURNAL
J Clin. Microbiol. 42 (9), 3925-3931 (2004)

PUBMED
15364970

REFERENCE
2 (bases 1 to 3357)

AUTHORS
La Scola, B., Khamis, A. and Raoult, D.
TITLE
Direct Submision
Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille
13385, France

FEATURES
source

1. .3357
Location/Qualifiers
/organism="Rhodococcus equi"
/mol_type="genomic DNA"
/strain="CIP 81.17"
/db_xref="taxon:43767"
/clone="retp056"
/note="conspecificity: Corynebacterium equi (type strain
CIP 54.72) = Corynebacterium hoagii (type strain CIP
81.17); type strain of Corynebacterium hoagii"
<1. .3357
/gene="rpoB"
<1. .3357
/gene="rpoB"
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="rpoB"
/protein_id="AAS89247.1"
/db_xref="GI:46361028"
/translation="SRQTKAVAGIIPGAPRRVSPAKIREPLEVPGLLDIQTSFENVG
SPWREDAARAGBGRVTSIGLEIDLELSPIDYSGNMSLSISPRREDYKTTIDEC
KDMTYAALFVTAEPINNDQELKSQTVFICGDFPLMTDKGTFVINGTERVVSOLVRS
PGVYFDSIDSTKTLHSVAVIIPSRGAMLEFDVKDRDVTGVRIDRRKORPEOETPRDIAOSIL

LGMTDEIRERPFGESEIMATLEKDPKANTDEALDITRKLRGPPEPTKSAQTLLN
LPEKDKRYDLARVGRYKINKKLGINTGLPIEASTLTEDDITVITIELVRLTAGDTMT
APGVEPVVDIDHFGNRLRLRTVGLIQLIIVGLSMEVREVRRETTDVAITP
QTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALGPGSLSERGLATP
RDVPSHYGMCPIETEPENIGLISLIVARVNPGRPIETPRKXENGLTPOUY
LITDEBRHVAQANSPYDANGRTERTPRVLRKGEVEPSSDVIMYDPSPMQMS
VAAAMIPELHDDANRALMGANNORAPVLRSEAPLVGTMEILAAVADADVITEK
TGAVEVSSDYVTVMADGSRRTYRLKRFANSNQTANORPIVDEQORVAGVAD
GPTCENEMALGKULIIVAMPBGNVEDAILISQAVEEDVITSIHIEHEIDART
KLGAEITRDIPIVNSDEVADLDGERGIVRIGAEVARDGILVAGTYPKGETLPEERL
LRAIFGKAREBDTSLKVPHGESGKIVIGRVREREDDLPDGVNELVRYVQKX
IQGDKLHAGHNGKVGITKILPQEDMPELPDGTVDITLINTHGVPRMNIGQVAFETH
GMTKGTGMVQIADGSRPDMATLPEEMLSAPDSNATATPEAKEDDELGLST
LPRNDEGRWVPPDKATLFDGSGEPPEPVSVGMVILKHLVDDKIHARSSTGPS
MITOOLPGKAGOFGGRFGMECWMQAYGAAYTLOEHLITKSDVDVGRVAYEAIYK
GENIPERGISFVILKELQSLCLNVEVL"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3357;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCG 16
|||||

Db 879 TGTAGCGGCCACGCG 864

RESULT 12

AY62743/c 3702 bp DNA linear BCT 13-JAN-2004
LOCUS Mycobacterium wolinskyi strain ATCC 700010 RpoB gene, complete cds.

AY62743
AY62743.1 GI:34595758

DEFINITION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

Adekanbi, T., Drancourt, M. and Raoult, D.

Submitted (26-MAR-2003) Unite des Rickettsies, Faculte de Medecine,

27 Boulevard Jean Moulin, Marseille 13385, France

Location/Qualifiers

1. 3702

/organism="Mycobacterium wolinskyi"

/mol_type="genomic DNA"

/strain="ATCC 700010"

/db_xref="taxon:700010"

/db_xref="taxon:59750"

/note="type strain of Mycobacterium wolinskyi"

193..3702

/codon_start=1

/product="RpoB"

/protein_id="AA076602.1"

/db_xref="GI:34595759"

/translation="MLEGCIILVSSGSSNAITNNSVGPANRVSFAKLREPLFVPLG
LDVOTSPFVLVSGDRWROAIDRGEPNPGGLEEVLAELSPIIDFSGMSLSDPR
FDEYKASVDECKDKMTYAPLFTYAEINNTSEIKSQVFMGDFPMWTKGFIIN
GTERVAVSOLVRSRGVYFDETIKSTETKLSVAVIPRGAMLEFDVAKRTVGRV
RKROPTVILKALGMNEQIVFRFSGEIMMGKXDTSGDEALDITRKLRPGE
PTKESQTLLENLFFKERYDLARVGRYKINKKLGINTGLPIEASTLTEDDITV
YLRHGGTMTVPGVEVVDIDHFGNRLRLRTVGLIQLIIVGLSMEVREVRRE
ERMTOVEAITPOTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALG
PGSLSERGLATPQTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALG
PGSLSERGLATPQTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALG
KVDVGVTDQIDYLTADBEDRHVAQANSPTDENGRTERTPRVLRKGEVEPSSDVIMYDPSPMQMS
VAAAMIPELHDDANRALMGANNORAPVLRSEAPLVGTMEILAAVADADVITEK
TGAVEVSSDYVTVMADGSRRTYRLKRFANSNQTANORPIVDEQORVAGVAD
GPTCENEMALGKULIIVAMPBGNVEDAILISQAVEEDVITSIHIEHEIDART
KLGAEITRDIPIVNSDEVADLDGERGIVRIGAEVARDGILVAGTYPKGETLPEERL
LRAIFGKAREBDTSLKVPHGESGKIVIGRVREREDDLPDGVNELVRYVQKX
IQGDKLHAGHNGKVGITKILPQEDMPELPDGTVDITLINTHGVPRMNIGQVAFETH
GMTKGTGMVQIADGSRPDMATLPEEMLSAPDSNATATPEAKEDDELGLST
LPRNDEGRWVPPDKATLFDGSGEPPEPVSVGMVILKHLVDDKIHARSSTGPS
MITOOLPGKAGOFGGRFGMECWMQAYGAAYTLOEHLITKSDVDVGRVAYEAIYK
GENIPERGISFVILKELQSLCLNVEVL"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3702;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCG 16
|||||

Db 1108 TGTAGCGGCCACGCG 1093

RESULT 13

AY147167/c 3868 bp DNA linear BCT 01-SEP-2003
LOCUS Mycobacterium septicum strain ATCC 700731T RpoB (rpoB) gene,
complete cds.

AY147167
AY147167.1 GI:34391552

DEFINITION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

Adekanbi, T., Drancourt, M. and Raoult, D.

Submitted (03-SEP-2002) Unite des Rickettsies, Faculte de Medecine,

Boulevard Jean Moulin, Marseille 13285, France

Location/Qualifiers

1. 3868

/organism="Mycobacterium septicum"

/mol_type="genomic DNA"

/strain="ATCC 700731T"

/db_xref="taxon:700731T"

/db_xref="taxon:98668"

/gene="rpoB"

118..3675

/gene="rpoB"

/codon_start=1

/product="RpoB"

/protein_id="AA065354.1"

/db_xref="GI:34391553"

/translation="MAANGPAYQLSGKSHVEFGCLAVSSGSTANANTNNSVPGAPN
RVSEFAKLEPLVGLDVTQDSFDMVAGDGRQAVDVGEGNPGGLEVELELSP
IEDPSGMSLSPDPRFDEYKASVDECKDKMTYAPLFTYAEINNTSEIKSQVFMGDFPMWTKGFIIN
GTERVAVSOLVRSRGVYFDETIKSTETKLSVAVIPRGAMLEFDVAKRTVGRV
RKROPTVILKALGMNEQIVFRFSGEIMMGKXDTSGDEALDITRKLRPGE
PTKESQTLLENLFFKERYDLARVGRYKINKKLGINTGLPIEASTLTEDDITV
YLRHGGTMTVPGVEVVDIDHFGNRLRLRTVGLIQLIIVGLSMEVREVRRE
ERMTOVEAITPOTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALG
PGSLSERGLATPQTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALG
PGSLSERGLATPQTLINIRPVAAIKPEFRTSOLQFOMDONNLSGLTHKRRLSALG
KVDVGVTDQIDYLTADBEDRHVAQANSPTDENGRTERTPRVLRKGEVEPSSDVIMYDPSPMQMS
VAAAMIPELHDDANRALMGANNORAPVLRSEAPLVGTMEILAAVADADVITEK
TGAVEVSSDYVTVMADGSRRTYRLKRFANSNQTANORPIVDEQORVAGVAD
GPTCENEMALGKULIIVAMPBGNVEDAILISQAVEEDVITSIHIEHEIDART
KLGAEITRDIPIVNSDEVADLDGERGIVRIGAEVARDGILVAGTYPKGETLPEERL
LRAIFGKAREBDTSLKVPHGESGKIVIGRVREREDDLPDGVNELVRYVQKX
IQGDKLHAGHNGKVGITKILPQEDMPELPDGTVDITLINTHGVPRMNIGQVAFETH
GMTKGTGMVQIADGSRPDMATLPEEMLSAPDSNATATPEAKEDDELGLST
LPRNDEGRWVPPDKATLFDGSGEPPEPVSVGMVILKHLVDDKIHARSSTGPS
MITOOLPGKAGOFGGRFGMECWMQAYGAAYTLOEHLITKSDVDVGRVAYEAIYK
GENIPERGISFVILKELQSLCLNVEVL"

/ tinnala = "mæskshvælgicllansvosskannitinnvgaupstærakr
 eflævþpðlvdytdsþmclvæmgöckvæðreðþkglævæelþpððsgsm
 slsþpðpððvkrþvðckðkmttaþlþftaþeinnntseikstþvngðþmmt
 þkgrftinngerwvavslanspæyþeðstðkxtællaviligrakællævðvkr
 dþvgridrkrævavtllakmtæmðltergæfþlmgællvðstæðallæd
 yrkrægrþrksaotllntþfkrærvðavægvçrkrkglænaoþltsstlnt
 edvavtællævrlægtlntþpægvævþvævndmhþænkrlævælltõnrlvæ
 smævæværrtðvæltipõlntinrþvvaikæfþetsoqlsþmõgnlsglþtæ
 krslslæpçglæsergælværvðshværmçlþtþpçgnlþlsglsværnæv
 gælrþvkrvçvgtvðtõllytlæðððhvvavvaoþlþpðgrþlævrvægv
 venvaþsvðvmdvþpõmvsatvntþlæðððnarlalmgæmõrøvlyvæslþ
 vçtgmællæddadþvæsekaþvævsvælytlvnaððsgstþlrmkrþvashnçt
 anõrþivagaovævavvðgþctægnællakmlvaþmæhçvæðlþlšnrl
 vævðvltstihæhððarõtkæsttðþvnsðvællððrgvrlvgaevrðð
 vðlþvævþkçkætlþbærlrlrlþfækrærvævrtslçpþvsgçlçrþvæv
 dðlþpægnævælvævvaõkrlsgððlærhænkçvlgkllþvðmþlþpçgvvð

[illegible]

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3929;

Best Local Similarity 100.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGGCCACGCG 16

Db 1108 TGTAGCGGGCCACGCG 1093

Search completed: February 10, 2005, 13:13:45
Job time : 120.196 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 19.6391 Seconds
(without alignments)
5727.097 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19

Sequence: 1 Tgtagcgccacgcgctg 19

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0	19	6	ABK15246	Abk15246 Corn tran
2	100.0	498	6	ABK15244	Abk15244 Corn nk60
3	84.2	208	4	AAS05219	Aas05219 Mycobacte
4	84.2	35167	12	ADMA43175	Adma43175 Wild-type
5	84.2	35167	13	ADS00140	Ads00140 Wild-type
6	78.9	752	4	AA159900	Aa159900 Human pol
7	78.9	1098	12	AD030319	Ado30319 Mouse GPC
8	78.9	1183	12	ADP72493	Adp72493 Renal tox
9	78.9	1253	10	ADBS2577	Adbs2577 Primary r
10	78.9	1405	2	AAQ46125	Aaq46125 PGE2 rece
11	78.9	1780	10	ABV75862	Abv75862 Human pot
12	78.9	2107	2	AAQ46124	Aaq46124 PGE2 rece
13	78.9	14800	10	ABV75863	Abv75863 Human pot
14	78.9	51855	10	ACF04818	Acf04818 Melithiaz
15	78.9	214520	10	ADL13471	Adl13471 Osteoarth
16	73.7	100	8	ACD72269	Acd72269 E. coli K
17	73.7	228	8	ABX54979	Abx54979 Bovine ES
18	73.7	228	13	ADBS5656	Adbs5656 Bacterial
19	73.7	292	8	ABX42572	Abx42572 Bovine ES
20	73.7	414	9	ACH29799	Ach29799 Human tes

ALIGNMENTS

21	14	73.7	436	8	ABX52378	Abx52378 Bovine ES
22	14	73.7	465	12	ACH92095	Ach92095 Human gen
23	14	73.7	500	4	AAD10051	Aad10051 Mouse 5'
24	14	73.7	619	13	ADR65184	Adr65184 Cotton cD
25	14	73.7	756	12	ADK13721	Adk13721 E. coli i
26	14	73.7	834	12	ADJ39824	Adj39824 Plant cDN
27	14	73.7	862	5	AAS74970	Aas74970 DNA encod
28	14	73.7	870	12	ADJ43711	Adj43711 Plant cDN
29	14	73.7	1004	5	AAS74969	Aas74969 DNA encod
30	14	73.7	1146	8	ACA26804	Ac26804 Prokaryot
31	14	73.7	1209	5	AB198018	Ab198018 Non-endog
32	14	73.7	1209	8	ABV76938	Abv76938 Nucleotid
33	14	73.7	1209	12	AD030027	Ado30027 Human GPC
34	14	73.7	1226	3	AA293894	Aa293894 Human EP-
35	14	73.7	1226	3	AAA34920	Aa34920 Human ade
36	14	73.7	1376	3	AA21042	Aa21042 Human low
37	14	73.7	1376	8	AB234821	Ab234821 Coding be
38	14	73.7	1376	8	AB242749	Ab242749 Human pro
39	14	73.7	1376	10	AB296736	Ab296736 Human nuc
40	14	73.7	1376	10	ACA56763	Aca56763 Human s1g
41	14	73.7	1376	11	ABD20585	Abd20585 Human pul
42	14	73.7	1376	12	AD156559	Ad156559 Human pol
43	14	73.7	1394	13	ADR46583	Adr46583 Human pro
44	14	73.7	1394	2	AAQ80287	Aaq80287 Prostagla
45	14	73.7	1431	8	ACA23808	Aca23808 Prokaryot

RESULT 1

ABK15246 standard; DNA; 19 BP.

ABK15246;

08-MAY-2002 (first entry)

Corn transgene junction probe #1.

Probe; ss; rice actin 1 promoter; RAI; RAI intron;

chloroplast transit peptide gene; glyphosate resistance; corn;

5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;

transcriptional terminator; cauliflower mosaic virus 35S promoter;

PV-ZMGT32; transgenic; nk603.

Zea mays.

Synthetic.

EP1167531-A1.

15-JUN-2001; 2001EP-00202314.

22-JUN-2000; 2000US-0213567P.

13-OCT-2000; 2000US-0240014P.

13-OCT-2000; 2000US-0241215P.

(MONS) MONSANTO TECHNOLOGY LLC.

Behr CF, Hironaka C, Heck GR, You J;

WPI; 2002-165871/22.

Novel DNA construct useful for producing a corn plant that tolerates

application of glyphosate herbicide, comprises two transgene expression

cassettes.

Claim 8; Page 3; 25pp; English.

The invention relates to a DNA construct (1) comprising a first (F1) and

a second (F2) expression cassette. F1 of the DNA construct, in operable

CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
CC peptidic gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator
CC (IV), and P2 has cauliflower mosaic virus 35S promoter, Hep70 intron,
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant
CC PV-ZMG732(hk603) of 498 or 1183bp, or primers and probes derived from
CC them used to detect transgene junction points). The construct is useful
CC for producing a corn plant that tolerates application of glyphosate
CC herbicide by transforming a corn cell with, selecting the corn cell for
CC tolerance to application of glyphosate, and growing the corn cell into a
CC fertile corn plant. The present sequence is a transgene junction probe
CC sequence consisting of corn genomic DNA and vector sequences
XX
SQ Sequence 19 BP; 2 A; 6 C; 8 G; 3 T; 0 U; 0 Other;

Query Match	100.0%	Score 19	DB 6	Length 19
Best Local Similarity	100.0%	Pred. NO. 0.14		
Matches 19, Conservative	0	Mismatches	0	Gaps 0

OY	1 TGTAGCGGCCACGCGTGG 19
D_B	1 TGTAGCGGCCACGCGTGG 19

RESULT 2
ABK15244
ID ABK15244 standard; DNA; 498 BP.

AC	ABK15244;
XX	
DT	08-MAY-2002 (first entry)

Corn nk603 5' transgene/genomic sequence PCR fragment.

KW ds; rice actin 1 promoter; RA1; RA1 intron;
KW chloroplast transit peptide gene; glyphosate resistance; corn;
KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSS; Hsp70 intron;
KW transcriptional terminator; cauliflower mosaic virus 35S promoter;
KW PV-ZMG132; transgenic; nk603.

OS Zea mays.
OS Oryza sativa.
OS Synthetic.
OS Chimeric.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
FT misc_feature 1..304
FT /*tag= a
FT /note= "Corn genomic sequence"
```

```
FT /note= "Vector sequence"
FT misc_feature 350, .498
FT /tag= C
FT /note= "Rice ACT1 sequence"
```

PN EP1167531-A1.

PD 02-JAN-2002.

PF 15-JUN-2001; 2001EP-00202314.

PR 22-JUN-2000; 2000US-0213567P.

PR 13-OCT-2000; 2000US-0241215P.

PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX

PI Behr CF, Hironaka C, Heck GR, You J,

DR WPI; 2002-165871/22.

PT . Novel DNA construct useful for producing a corn plant that tolerates

PT application of glyphosate herbicide, comprises two transgene expression
PT cassettes.
XX
PS Claim 3, Page 16, 25pp; English.

CC The invention relates to a DNA construct (I) comprising a first (P1) and
CC a second (P2) expression cassette. P1 of the DNA construct, in operable
CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator
CC (IV), and P2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant
CC pV-ZMG732(hk603) of 498 or 1183bp, or primers and probes derived from
CC them (used to detect transgene junction points). The construct is useful
CC for producing a corn plant that tolerates application of glyphosate
CC herbicide by transforming a corn cell with, selecting the corn cell for
CC tolerance to application of glyphosate, and growing the corn cell into a
CC fertile corn plant. The present sequence is the 48bp amplicon consisting
CC of corn genomic sequence, vector sequences and Rice ACT1 promoter
CC sequences

Sequence 498 BP; 148 A; 114 C; 131 G; 105 T; 0 U; 0 Other;

Query Match	100.0%;	Score 19;	DB 6;	Length 498;
Best Local Similarity	100.0%;	Pred. No. 0.097;		
Matches	19;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

[illegible]

RESULT 3
AAS05219/c
ID AAS05219 standard; DNA; 208 BP.

AC AAS05219;

DT 07-SEP-2001 (first entry)

DE *Mycobacterium fortuitum* rpoB gene fragment.

KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS *Mycobacterium fortuitum*.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR001223.

PR 27-OCT-1999; 99KR-00046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the *ryoB* gene of mycobacteria, useful for

PT Fragment length polymorphism.

PS Claim 1; Page 46; 50pp; English.

CC The present sequence for *Mycobacterium fortuitum* rpoB gene fragment is 1
CC of 24 rpoB gene fragments (AMS05201-AMS05224) from various *Mycobacteri*
CC species. These rpoB gene fragments are used in the diagnosis and
CC identification of *Mycobacterium* species using a novel PCR-restriction
CC fragment length polymorphism analysis (PRA) method. The method comprises
CC obtaining a restriction fragment length polymorphism (RFLP) pattern of

CC the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA
 CC fragment from the microorganism to be identified and comparing the RFLP
 CC patterns from the known rpoB gene fragments with the unidentified
 CC fragment. The rpoB gene fragments are useful to identify a wide range of
 CC Mycobacterium species, e.g. for diagnosis or to obtain epidemiological
 CC and pathogenesis information for selection of appropriate therapies
 CC including M. tuberculosis, M. lepre and non-tuberculous mycobacteria
 CC (NTM) encountered in subjects infected with human immunodeficiency virus
 CC (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and
 CC cost effective (only 1 PCR required), and can differentiate between many
 CC species in a single experiment, including those difficult to distinguish
 CC by usual biochemical tests. Also described are oligonucleotide probes
 CC (AAS05227-AAS05242) for detecting specific Mycobacterial species
 CC
 SQ Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 4; Length 208;
 Best Local Similarity 100.0%; Pred. No. 6.1; Mismatches 0; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGGCCCGCG 16
 |||||
 Db 42 TGTAGCGGCCCGCG 27

RESULT 4
 ID ADM43175/c
 ADM43175; standard; DNA; 35167 BP.

AC ADM43175;
 DT 03-JUN-2004 (first entry)

XX Wild-type adenovirus serotype 24 nucleotide sequence SEQ ID NO:5.

XX replication-defective adenovirus; adenoviral E1-complementing cell line;
 KM anti-HIV; gene therapy; vaccine; immune response; HIV;

XX adenovirus serotype 24; gene; ds.

XX unidentified adenovirus.

OS WC2004018627-A2.

XX 04-MAR-2004.

PD 21-AUG-2003; 2003WO-US026145.

XX 22-AUG-2002; 2002US-0405182P.

PR 17-MAR-2003; 2003US-0455234P.

PR 17-MAR-2003; 2003US-0455312P.

PR 28-MAR-2003; 2003US-0458825P.

XX (MERI) MERCK & CO INC.

PI Belt AJ, Chastain M, Sandig V, Emin EA, Shiver JW, Casimiro DR;
 PI Kaslow DC, Morsy M;

XX WPI; 2004-226820/21.

DR Propagating replication-defective adenovirus in an adenoviral E1-
 XX complementing cell line, useful in gene therapy applications, comprises
 PT providing all or a portion of an E4 region in cis within the genome of
 PT the adenovirus.

XX Example 7; SEQ ID NO 5; 113pp; English.

XX The present invention describes a method for propagating a replication-
 CC defective adenovirus in an adenoviral E1-complementing cell line
 CC expressing E1 gene product(s) which are non-native to the adenovirus,
 CC comprising providing all or a portion of an E4 region in cis within the
 CC genome of the replication-defective adenovirus. Also described: (1) a
 CC replication-defective adenovirus that comprises all or a portion of a
 CC heterologous E4 region comprising a heterologous adenoviral ORF, or that

CC is propagated by the method described above; (2) a recombinant adenoviral
 CC vector of serotype 24 which comprises an E4 gene or a segment of an E4
 CC gene comprising ORF6 of an alternative serotype; (3) a population of
 CC cells comprising the recombinant adenoviral vector; (4) producing
 CC recombinant, replication-defective adenovirus particles; (5) a purified
 CC recombinant, replication-defective adenovirus particles harvested in
 CC accordance with the method in (4); (6) a composition comprising purified
 CC recombinant adenovirus particles; (7) effecting the delivery and
 CC expression of heterologous nucleic acid, and (8) generating a cellular-
 CC mediated immune response against HIV in an individual. The replication-
 CC defective adenovirus has anti-HIV activity, and can be used in gene
 CC therapy and in vaccines. The method can be used for propagating a
 CC replication-defective adenovirus in an adenoviral E1-complementing cell
 CC line useful for propagating and rescuing multiple serotypes of
 CC replication-defective adenovirus in a single adenoviral E1-complementing
 CC cell line and for generating a cellular-mediated immune response against
 CC HIV in an individual. The resultant virus can be studied and used in
 CC various gene therapy and vaccine efforts. The present sequence represents
 CC a wild-type adenovirus serotype 24 nucleotide sequence, which is used in
 CC the exemplification of the present invention.

SQ Sequence 35167 BP; 7977 A; 9958 C; 9996 G; 7236 T; 0 U; 0 Other;
 Query Match 84.2%; Score 16; DB 12; Length 35167;
 Best Local Similarity 100.0%; Pred. No. 3.4; Mismatches 0; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGGCCCGCG 16
 |||||
 Db 12091 TGTAGCGGCCCGCG 12076

RESULT 5
 ADS00140/c
 ID ADS00140 standard; DNA; 35167 BP.

XX ADS00140;

AC 16-DEC-2004 (first entry)

XX Wild-type human adenovirus serotype 24 DNA SEQ ID NO:1.

XX ds; adenovirus serotype 24; anti-HIV; vector; gene therapy; vaccine.

XX Human adenovirus type 24.

XX WC2004083418-A1.

XX 30-SEP-2004.

XX 21-AUG-2003; 2003WO-US026338.

XX 17-MAR-2003; 2003US-0455312P.

XX (MERI) MERCK & CO INC.

PI Emin EA, Shiver JW, Belt AJ, Casimiro DR, Chastain M, Kaslow DC;
 PI Morsy MA;

XX WPI; 2004-691052/67.

DR New recombinant adenoviral vector of serotype 24 which is partially
 XX deleted in E1 and devoid of E1 activity, useful for delivering and
 PT expressing exogenous genes, e.g. in gene therapy or in vaccination
 PT protocols.

XX Disclosure; SEQ ID NO 1; 133pp; English.

XX The invention relates to a novel recombinant adenoviral vector of
 CC serotype 24 which is partially deleted in E1 and devoid of E1 activity
 CC and comprises a heterologous nucleic acid or an human immunodeficiency
 CC virus (HIV)-1 gene. A vector of the invention has anti-HIV activity, and
 CC may have a use in gene therapy, and as a vaccine. The recombinant

CC adenoviral vector of serotype 24 which is at least partially deleted in
CC E1 and devoid of E1 activity is useful for the delivery and expression of
CC exogenous genes. The adenoviral vectors are useful in gene therapy or in
CC vaccination protocols. The present sequence represents the nucleic acid
CC sequence for wild-type adenovirus serotype 24.

XX Sequence 35167 BP; 7977 A; 9958 C; 9996 G; 7236 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 13; Length 35167;
Best Local Similarity 100.0%; Pred. No. 3,4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16
DB 12091 TGTAGCGGCCCGCG 12076

RESULT 6
AA159900/c
ID AA159900 standard; cDNA; 752 BP.

XX AC AA159900;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 3889.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukemia; ss.

XX Homo sapiens.

XX WO20015312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-0052317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX PI Zhou P, Goodrich R, Dirmann RT;
XX DR WPI: 2001-442253/47.
XX DR P-PSDB; AAM40744.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.

XX PS Claim 1; SEQ ID NO 3889; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AAM38642-AA42213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemoreactive/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification

XX Sequence 752 BP; 184 A; 156 C; 233 G; 179 T; 0 U; 0 Other;

Query Match 78.9%; Score 15; DB 4; Length 752;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCCGCGT 17
DB 30 TAGCGGCCCGCGT 16

RESULT 7
AD030319/c
ID AD030319 standard; cDNA; 1098 BP.

XX AC AD030319;

XX DT 29-JUL-2004 (first entry)

XX DE Mouse GPCR PIGR3 polynucleotide, SEQ ID NO:1422.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
XX transgenic mouse; neurological disorder; adrenal gland disorder;
XX colon disorder; intestinal disorder; cardiovascular disorder;
XX muscular disorder; blood disorder; immune disorder; bone disorder;
XX joint disorder; metabolic disorder; nutritive disorder; cancer;
XX kidney disorder; liver disorder; lung disorder; breast disorder;
XX ovary disorder; uterus disorder; prostate disorder; testis disorder;
XX skin disorder; stomach disorder; pancreas disorder; spleen disorder;
XX thymus disorder; thyroid disorder; antiparkinsonian; autimanic;
XX cyostatic; antiinflammatory; vasotropic; antidiarrhoeal; antidiabetic;
XX CNS; central nervous system; respiratory; antidiarrhoeal; antidiabetic;
XX vitruclide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;
XX dermatological; antitumor; antithyroid; antiallergic; anorectic;
XX immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;
XX murine; gene; ss.

XX OS Mus musculus.

XX WO2004040000-A2.

XX PD 13-MAY-2004.

XX PF 09-SEP-2003; 2003WO-US028226.

XX PR 09-SEP-2002; 2002US-0409303P.
XX PR 09-APR-2003; 2003US-0461329P.

XX PA (PRIM-) PRIMAL INC.

XX PI Galtanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
XX PI Madisen L, McIlwain KU, Pavlova KM, Vasiliadis D, Zeng H;
XX DR WPI: 2004-390329/36.
XX DR P-PSDB; ADO29621.

XX PT Novel mammalian G protein coupled receptors, useful for identifying
XX compounds that modulates diagnosing and treating disease condition
XX associated with GPCR dysfunction e.g. autoimmune diseases, angina
XX pectoris, Parkinson's disease.

XX Claim 151; SEQ ID NO 1422; 542bp; English.

XX The invention relates to human and mouse G protein-coupled receptors

(GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several mice, each of which has a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridise to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia); disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and thyroid (e.g., cancers). The present sequence represents a GPCR-encoding nucleic acid of the invention. Note: The full sequence data for this patent did not form part of the printed specification; those sequences not shown were obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences.

Sequence 1098 BP; 185 A; 342 C; 335 G; 236 T; 0 U; 0 Other;
Query Match 78.9%; Score 15; DB 12; Length 1098;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGC 15
Db 526 TGTAGCGGCCACGC 512

RESULT 8
ADP72493/c
ID ADP72493 standard; DNA; 1183 BP.

AC ADP72493;

DT 26-AUG-2004 (first entry)

DE Renal toxin progression gene marker #1082.

XX ds: toxic effect; gene expression profile; kidney tissue;
KM differential gene expression; toxicity progression; toxicity marker;
KM drug screening; toxicity assay; kidney pathology; nephritis;
KM kidney necrosis; glomerular injury; tubular injury;
KM focal segmental glomerulosclerosis.

XX Rattus norvegicus.

OS WO2004048598-A2.

PN 10-JUN-2004.

PD 24-NOV-2003; 2003WO-US037556.

PR 22-NOV-2002; 2002US-00301856.

XX (GENE-) GENE LOGIC INC.

PA Mendrick DL, Porter MM, Johnson KR, Castle A, Higgs B;

PI Elsholt M;

DR WPI; 2004-460771/43.

XX predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.

PS Claim 11; SEQ ID NO 1082; 266pp; English.

XX The invention relates to a method of predicting (the progression of) a
XX toxic effect of a compound by preparing a gene expression profile of a
XX kidney tissue or cell sample exposed to the compound and comparing the
XX gene expression profile to a database, or detecting the level of gene(s)
XX expression in a tissue or cell sample exposed to the compound, where
XX differential gene expression compared to a control indicates a toxic
XX effect (toxicity progression). The method is useful for predicting (the
XX progression of) at least one toxic effect of a compound. The genes are
XX useful as toxicity markers in drug screening and toxicity assays. The
XX methods are useful for predicting the likelihood that a compound or test
XX agent will induce various specific kidney pathologies, such as nephritis,
XX kidney necrosis, glomerular and tubular injury, or focal segmental
XX glomerulosclerosis. The methods are useful for determining the similarity
XX of a toxic response to one or more individual compounds and for
XX predicting or elucidating the potential cellular pathways influenced,
XX induced or modulated by the compound or test agent. The kit is useful for
XX predicting or modelling the toxic response of a test compound, for
XX monitoring the progression of renal disease states, for identifying genes
XX that show promise as new drug targets and for screening known and newly
XX designed drugs. This sequence corresponds to a gene marker used in the
XX method of the invention. (Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences).

Sequence 1183 BP; 208 A; 361 C; 366 G; 248 T; 0 U; 0 Other;
Query Match 78.9%; Score 15; DB 12; Length 1183;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGC 15
Db 582 TGTAGCGGCCACGC 568

RESULT 9
ADB52577/c
ID ADB52577 standard; DNA; 1253 BP.

AC ADB52577;

DT 04-DEC-2003 (first entry)

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3119.

XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KM primary rat hepatocyte toxicity modelling; gene; ds.

XX Rattus norvegicus.

OS WO2003065993-A2.

PN 14-AUG-2003.

PD 04-FEB-2003; 2003WO-US003482.

PR 04-FEB-2002; 2002US-0353171P.

PR 13-MAR-2002; 2002US-0363539P.

PR 08-APR-2002; 2002US-0370246P.

PR 10-APR-2002; 2002US-0371134P.

PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.

PR 11-APR-2002; 2002US-0371413P.

PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of tox mean and non-tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 3119; 874bp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the tox mean and non-tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX
SQ Sequence 1253 BP; 213 A; 386 C; 373 G; 281 T; 0 U; 0 Other;
Query Match 78.9%; Score 15; DB 10; Length 1253;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTAGCGGCCACGC 15
DB 614 TGTAGCGGCCACGC 600
RESULT 10
AAQ46125/c
ID AAQ46125 standard; cDNA; 1405 BP.
XX
XX AAQ46125;
XX
XX 25-MAR-2003 (revised)
DT 09-FEB-1994 (first entry)
XX
XX PGE2 receptor (EP3 beta) clone MP653.
XX
XX prostaglandin E receptor; PGE; agonist; antagonist; GTP-binding protein;
KM digestive tract; constriction; relaxation; gastric acid;
KM intestinal juice; neurotransmitter; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
FH 106..1191
FT /*tag= a
FT /product= "PGE2_receptor"
FT

FT /note= "PGE subtype EP3 beta"
XX
XX EP557966-A1.
XX
XX 01-SEP-1993.
PD
XX 24-FEB-1993; 93EP-00102873.
XX
XX 24-FEB-1992; 92JP-00036580.
PR 23-MAR-1992; 92JP-00064889.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Ichikawa A, Narumiya S;
PI
XX WPI; 1993-274435/35.
DR P-PDB; AAR42281.
XX
XX New prostaglandin E receptor protein and DNA encoding it - used to study
PT prostaglandin binding, agonists and antagonists etc.
XX
XX Claim 15; Fig 4; 22bp; English.
XX
XX PGE2 is known to be involved in digestive tract constriction/relaxation,
CC gastric juice secretion and neurotransmitter release. The protein encoded
CC by the gene is capable of receiving PGE and of activating GTP-binding
CC protein. MP653 (AAQ46125) is identical to MP660 (AAQ46124) with the
CC exception of a 89-bp deletion in the coding region of the C-terminal tail
CC of the receptor. This deletion creates a new reading frame downstream
CC from this junction. This results in a 30 amino acid C-terminal fragment
CC of the receptor being replaced with a new 26-amino acid fragment in the C
CC -terminal end of the MP653 (AAQ46125) receptor. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
SQ Sequence 1405 BP; 251 A; 415 C; 418 G; 321 T; 0 U; 0 Other;
Query Match 78.9%; Score 15; DB 2; Length 1405;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTAGCGGCCACGC 15
DB 631 TGTAGCGGCCACGC 617
RESULT 11
ABV75862/c
ID ABV75862 standard; cDNA; 1780 BP.
XX
XX ABV75862;
XX
XX 05-FEB-2003 (first entry)
DT
XX
XX Human potassium channel transporter cDNA.
XX
XX Transporter; potassium channel; human; gene therapy; chromosome 17; gene;
XX ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..374
FT /*tag= a
FT 375..1667
FT CDS /*tag= b
FT /product= "Transporter protein"
FT 1667..1780
FT /*tag= c
XX
XX WO200281657-A2.
XX
XX 17-OCT-2002.
XX

PF 01-APR-2002; 2002WO-US009743.
 XX 30-MAR-2001; 2001US-00820923.
 XX (PEKE) PE CORP NY.
 PA (BEAS/) BEASLEY E M.
 PA (WEIM/) WEI M.
 PA (KETC/) KETCHUM K A.
 PA (DFRA/) DI FRANCESCO V.
 XX
 PI Beasley EM, Wei M, Ketchum KA, Di Francesco V;
 XX
 DR WPI; 2003-058533/05.
 DR P-PSDB; ABP55112.
 XX
 PT Novel isolated human transporter peptide useful for treating disorder
 PT characterized by absence of, in appropriate or unwanted expression of the
 PT transporter protein, and as immunogens to raise antibodies.
 XX
 PS Claim 4; Fig 1; 75pp; English.
 XX
 CC The present sequence is that of cDNA encoding a novel human transporter
 CC protein that is related to the voltage gated potassium channel
 CC transporter family, and which is expressed in the lung and in the foetal
 CC brain. The gene (see ABV75863) encoding the transporter is located on
 CC chromosome 17. Nucleic acids encoding the transporter are useful as
 CC probes, primers, in biological assays, for constructing recombinant
 CC vectors, host cells and transgenic animals, for expressing antigenic
 CC portions of the human transporter protein, for designing ribozymes, for
 CC monitoring the effectiveness of modulating compounds on the expression or
 CC activity of the human transporter gene in clinical trials or in a
 CC treatment regimen, in diagnostic assays for qualitative changes in human
 CC transporter nucleic acid that lead to pathology, for testing an
 CC individual for a genotype that while not necessarily causing a disease,
 CC nevertheless affects the treatment modality, as antisense constructs to
 CC control human transporter gene expression in cells, tissues and
 CC organisms, and for gene therapy in patients containing cells that are
 CC aberrant in human transporter gene expression. A gene chip comprising a
 CC transporter nucleic acid is claimed, and is useful for conducting cell-
 CC based assays involving the human transporter protein
 XX
 XX Sequence 1780 BP; 352 A; 570 C; 496 G; 362 T; 0 U; 0 Other;
 SQ
 Query Match 78.9%; Score 15; DB 10; Length 1780;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 AGCGGCCCGCGGTG 18
 Db 51 AGCGGCCCGCGGTG 37
 RESULT 12
 AAQ46124/c
 ID AAQ46124 standard; cDNA; 2107 BP.
 XX
 AC AAQ46124;
 XX
 DT 25-MAR-2003 (revised)
 DT 09-FEB-1994 (first entry)
 XX
 DE PGE2 receptor (BP3 alpha) clone MP660.
 XX
 KW prostaglandin E receptor; PGE; agonist; antagonist; GTP-binding protein;
 KW digestive tract; constriction; relaxation; gastric acid;
 KW intestinal juice; neurotransmitter; ss.
 XX
 XX Mus musculus.
 OS
 XX
 FT Key Location/Qualifiers
 FT CDS 119..1216
 FT /tag= a
 FT /product= "PGE2_receptor"

FT /note= "PGE subtype BP3 alpha"
 XX
 XX BP57966-A1.
 XX
 XX
 PD 01-SEP-1993.
 XX
 PD 24-FEB-1993; 93EP-00102873.
 XX
 PP 24-FEB-1992; 92JP-00036580.
 XX
 PR 23-MAR-1992; 92UP-00064889.
 XX
 PA (TAKA) TAKEDA CHEM IND LTD.
 XX
 PI Ichikawa A, Naramiya S;
 XX
 DR WPI; 1993-274435/35.
 DR P-PSDB; AAR42280.
 XX
 PT New prostaglandin E receptor protein and DNA encoding it - used to study
 PT prostaglandin binding, agonists and antagonists etc.
 XX
 PS Claim 15; Fig 1; 22pp; English.
 XX
 CC PGE2 is known to be involved in digestive tract constriction/ relaxation,
 CC gastric juice secretion and neurotransmitter release. The protein encoded
 CC by the gene is capable of receiving PGE and of activating GTP-binding
 CC protein. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 2107 BP; 482 A; 549 C; 563 G; 513 T; 0 U; 0 Other;
 Query Match 78.9%; Score 15; DB 2; Length 2107;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTAGCGGCCCGCGC 15
 Db 644 TGTAGCGGCCCGCGC 630
 RESULT 13
 ABV75863/c
 ID ABV75863 standard; DNA; 1480 BP.
 XX
 AC ABV75863;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Human potassium channel transporter gene.
 XX
 KW Transporter; potassium channel; human; gene therapy; chromosome 17; gene;
 KW ds.
 XX
 XX Homo sapiens.
 OS
 XX
 FT Key Location/Qualifiers
 FT CDS 2374..12713
 FT /tag= a
 FT /product= "Transporter protein"
 FT /note= "contains introns"
 FT 2374..2449
 FT /tag= b
 FT /number= 1
 FT 2450..4293
 FT /tag= c
 FT /number= 1
 FT 4294..4527
 FT /tag= d
 FT /number= 2
 FT 4528..5092
 FT /tag= e
 FT /number= 2
 FT 5093..5220
 FT /tag= f
 FT exon
 FT intron
 FT intron
 FT exon


```

DE Osteoarthritis-associated polymorphic nucleotide #3.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX joint space narrowing; osteophyte development; joint pain;
XX osteoarthritis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
XX WO2003054166-A2.
XX
XX 03-JUL-2003.
XX
XX 19-DEC-2002; 2002WO-US041225.
XX PF
XX 20-DEC-2001; 2001US-0342603P.
XX PR
XX (INCY-) INCYTE GENOMICS INC.
XX PA
XX Jones KA, Schafer A;
XX PI
XX WPI; 2003-559141/52.
XX DR
XX
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX
XX Disclosure; SEQ ID NO 3; 297pp; English.
XX
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX the sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).
XX
XX Sequence 214520 BP; 57899 A; 52039 C; 51001 G; 53577 T; 0 U; 4 Other;
SQ
Query Match 78.9%; Score 15; DB 10; Length 214520;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTAGCGGGCCACGC 15
DB 120357 TGTAGCGGGCCACGC 120343

```

Search completed: February 10, 2005, 13:44:04
 Job time : 22.6391 secs

This Page Blank (uspto)

Qy	181	AGCGCGCGTGGCATTTCGGCBAACCTCTCTCTCGGATAGAGCCGATCTTTCTG	240
Db	181	AGCGCGCGTGGCATTTCGGCBAACCTCTCTCTCGGATAGAGCCGATCTTTCTG	240
Qy	241	GCAATTCCAAACCTTAGAGAGTGCCTCCCTGGTGGCTGCTCGGCCAGCAAGCCTTGTAG	300
Db	241	GCAATTTCAAACCTTAGAGAGTGCCTCCCTGGTGGCTGCTCGGCCAGCAAGCCTTGTAG	300
Qy	301	CGGCCCAAGCGGTGGTACCAAGCTTGATATCCCTAGAGGCGCGCGTTAACAAAGTTAATC	360
Db	301	CGGCCCAAGCGGTGGTACCAAGCTTGATATCCCTAGAGGCGCGCGTTAACAAAGTTAATC	360
Qy	361	GAGGTATTCTATATGCTTGAGAAAGAGAGTGGGATAGTCCAAATATATAACAAAGTAAAGA	420
Db	361	GAGGTATTCTATATGCTTGAGAAAGAGAGTGGGATAGTCCAAATATATAACAAAGTAAAGA	420
Qy	421	TTACCGGTCAAAAGTAAACATCAGTTAAAGGTATATTAAGTAAATATTCGGTATTA	480
Db	421	TTACCGGTCAAAAGTAAACATCAGTTAAAGGTATATTAAGTAAATATTCGGTATTA	480
Qy	481	AAGGTGGCCCAAGTGA	498
Db	481	AAGGTGGCCCAAGTGA	498

RESULT 2

US-09-872-051-8/c
Sequence 8, Application US/09872051
Patent No. 6825400
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZWGT32(nk603) and Composition and Methods for Detecting and Measuring Presence of the Same
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea maize plastid genes, tpell and tpoa
OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

	Query Match	23.7%	Score 118	DB 4	Length 1183
	Best Local Similarity	100.0%	Pred. No. 2.7e-50		
	Matches 118	Conservative	0	Mismatches 0	Indels 0
QY	308	CGCGTGTACCAAGCTTGATATTCCTTAGGGCGCGCGCTTAAACAAGCTTATCTCGAGGTCA	367		
Db	381	CGCGGTGTACCAAGCTTGATATTCCTTAGGGCGCGCGCTTAAACAAGCTTATCTCGAGGTCA	322		
QY	368	TTCATATCTCTTGAGAAGAGAGTGGGATATGTCACAAATATAAACAAAGGTATAGATTACC	425		
Db	321	TTCATATCTCTTGAGAAGAGAGTGGGATATGTCACAAATATAAACAAAGGTATAGATTACC	264		

RESULT 3
 US-09-441-340-27
 : Sequence 27, Application US/09441340
 : Patent No. 6448476
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Barry, Gerard F.
 : TITLE OF INVENTION: Phosphonate Metabolizing Plants
 : FILE REFERENCE: 38-21 (15303)
 : CURRENT APPLICATION NUMBER: US/09/441,340
 : CURRENT FILING DATE: 1999-11-16
 : EARLIER APPLICATION NUMBER: 60/108,763
 : EARLIER FILING DATE: 1998-11-17
 : NUMBER OF SEQ ID NOS: 32
 :
 : SOFTWARE: PatentIn Ver. 2.0
 :
 : SEQ ID NO 27
 :
 : LENGTH: 2378

FEAT

```

OTHER INFORMATION: Description of Artificial Sequence:expression
OTHER INFORMATION: cassette comprising a plant promoter linked to an
OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
OTHER INFORMATION: transferase, and a termination sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (28) .. (965)
FEATURE:
NAME/KEY: intron
LOCATION: (966) .. (1423)
FEATURE:
NAME/KEY: transit_peptide
LOCATION: (1440) .. (1667)
FEATURE:
NAME/KEY: CDS
LOCATION: (1668) .. (2099)
FEATURE:
NAME/KEY: terminator
LOCATION: (2114) .. (2369)
US-09-441-340-27

```

Query Match	20.3%	Score 101;	DB 3;	Length 2378;
Best Local Similarity	100.0%	Pred. No. 1.3e-41;		
Matches 101; Conservative	0;	Mismatches	0;	Gaps 0

Qy 325 GATATCCCTAGGGCGGCGCGTTAACCAACTTACTCAGAGTCATTCAATGCTTGAGAG 38

Db 1 GATATCCCTAGGGCGGCGCGTTAACCAACTTACTCAGAGTCATTCAATGCTTGAGAG 60

Qy 385 AGAGTCGGGATAGTCCAAAATAAAACAAAGTTAGATTACC 425

Db 61 AGAGTCGGGATAGTCCAAAATAAAACAAAGTTAGATTACC 101

RESULT 4

US-08-144-602B-7
: Sequence 7B, Application US/08144602B
: Patent No. 5641876
: GENERAL INFORMATION:
: APPLICANT: MCELROY, David
: APPLICANT: WU, Ray
: TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
: STREET: CLINTON SQUARE, P.O. BOX 1051
: CITY: ROCHESTER
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 14603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/144,602B
FILING DATE: 27-OCT-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/10140
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-144-602B-7

Query Match 14.1%; Score 70; DB 1; Length 1404;
Best Local Similarity 100.0%; Pred. No. 9.2e-26;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
DB 9 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 68

QY 416 TAAATTACC 425
DB 69 TAAATTACC 78

RESULT 5
US-08-144-602B-5
Sequence 5, Application US/08144602B
Patent No. 5641876
GENERAL INFORMATION:
APPLICANT: McElroy, David
TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: CLINTON SQUARE, P.O. BOX 1051
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/144,602B
FILING DATE: 27-OCT-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/10140
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2199 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-144-602B-5

Query Match 14.1%; Score 70; DB 1; Length 2199;
Best Local Similarity 100.0%; Pred. No. 9.1e-26;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
DB 806 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 865

QY 416 TAAATTACC 425
DB 866 TAAATTACC 875

RESULT 6
US-08-144-602B-4
Sequence 4, Application US/08144602B
Patent No. 5641876
GENERAL INFORMATION:
APPLICANT: McElroy, David
TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: CLINTON SQUARE, P.O. BOX 1051
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/144,602B
FILING DATE: 27-OCT-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/10140
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-144-602B-4

Query Match 14.1%; Score 70; DB 1; Length 5643;
Best Local Similarity 100.0%; Pred. No. 9e-26;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 415
DB 809 TACTCGAGTCATTGATGCTTGAGAGAGAGTCGGGATAGTCCAAATTAACAAGG 868

QY 416 TAAATTACC 425
DB 869 TAAATTACC 878

RESULT 7
US-08-144-602B-6

```
/ Sequence 6, Application US/08144602B
/ Patent No. 5641876
/ GENERAL INFORMATION:
/ APPLICANT: McElroy, David
/ TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
/ NUMBER OF SEQUENCES: 27
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
/ STREET: CLINTON SQUARE, P.O. BOX 1051
/ CITY: ROCHESTER
/ STATE: NEW YORK
/ COUNTRY: USA
/ ZIP: 14603
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/144,602B
/ FILING DATE: 27-OCT-1993
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: TIMIAN, SUSAN J.
/ REGISTRATION NUMBER: 34,103
/ REFERENCE/DOCKET NUMBER: 19603/10140
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 716-263-1636
/ TELEFAX: 716-263-1600
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1392 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-08-144-602B-6

Query Match      13.7%; Score 68; DB 1; Length 1392;
Best Local Similarity 100.0%; Pred. No. 9.7e-25;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGGTCATTGATGCTTGAGAGAGAGCTCGGATAGTCCAAATATTAACAAGGTA 417
DB      1 CTCGAGGTCATTGATGCTTGAGAGAGAGCTCGGATAGTCCAAATATTAACAAGGTA 60

QY      418 AGATTACC 425
DB      61 AGATTACC 68

RESULT 8
US-09-068-101-5
/ Sequence 5, Application US/09068101
/ Patent No. 6372960
/ GENERAL INFORMATION:
/ APPLICANT: PLANT GENETIC SYSTEMS N.V.
/ TITLE OF INVENTION: Improved Barstar Gene
/ FILE REFERENCE: 2121-139P
/ CURRENT APPLICATION NUMBER: US/09/068,101
/ EARLIER FILING DATE: 1998-08-26
/ EARLIER APPLICATION NUMBER: EP 96202446.9
/ EARLIER FILING DATE: 1996-09-03
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 4032
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: other nucleic
/ OTHER INFORMATION: acid, "Plasmid pmv71"
```

```
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1999)..(3400)
/ OTHER INFORMATION: label = PRAc1, "promoter region of rice actin gene"
/ OTHER INFORMATION: label = contains an intron in the leader"
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (3401)..(3676)
/ OTHER INFORMATION: label = barstar, "barstar DNA"
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (3677)..(4003)
/ OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
/ OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium"
/ OTHER INFORMATION: T-DNA"
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (3395)..(3404)
/ OTHER INFORMATION: label = NcoI, "NcoI recognition site"
/ OTHER INFORMATION: label = KpnI, "KpnI recognition site"
/ US-09-068-101-5

Query Match      13.5%; Score 67; DB 3; Length 4032;
Best Local Similarity 100.0%; Pred. No. 3.1e-24;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      359 TCGAGGTCATTGATGCTTGAGAGAGAGCTCGGATAGTCCAAATATTAACAAGGTA 418
DB      1999 TCGAGGTCATTGATGCTTGAGAGAGAGCTCGGATAGTCCAAATATTAACAAGGTA 2058

QY      419 GATTACC 425
DB      2059 GATTACC 2065

RESULT 9
US-09-970-921-5
/ Sequence 5, Application US/09970921
/ Patent No. 6759575
/ GENERAL INFORMATION:
/ APPLICANT: Frank Michiels et al.
/ TITLE OF INVENTION: Improved Barstar Gene
/ FILE REFERENCE: 2428-0108P
/ CURRENT APPLICATION NUMBER: US/09/970,921
/ EARLIER FILING DATE: 2001-10-05
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 4032
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: other nucleic
/ OTHER INFORMATION: acid, "Plasmid pmv71"
/ NAME/KEY: misc_feature
/ LOCATION: (1995)..(3400)
/ OTHER INFORMATION: label = PRAc1, "promoter region of rice actin gene"
/ OTHER INFORMATION: label = contains an intron in the leader"
/ NAME/KEY: misc_feature
/ LOCATION: (3401)..(3676)
/ OTHER INFORMATION: label = barstar, "barstar DNA"
/ NAME/KEY: misc_feature
/ LOCATION: (3677)..(4003)
/ OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
/ OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium"
/ OTHER INFORMATION: T-DNA"
/ NAME/KEY: misc_feature
/ LOCATION: (3395)..(3404)
/ OTHER INFORMATION: label = NcoI, "NcoI recognition site"
/ OTHER INFORMATION: label = NcoI, "NcoI recognition site"
```

```
; LOCATION: (4016)..(4021)
; OTHER INFORMATION: label = Kpnl, "Kpnl recognition site"
US-09-970-921-5

Query Match
Best Local Similarity 100.0%; Pred. No. 3.1e-24; Length 4032;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 TCGAGGTCATTCATATGCTTGAGAGAGAGTCCGAGATGTCCTCAAAATTAACCAAGGTAA 418
Db 1999 TCGAGGTCATTCATATGCTTGAGAGAGAGTCCGAGATGTCCTCAAAATTAACCAAGGTAA 2058

QY 419 GATTACC 425
Db 2059 GATTACC 2065

RESULT 10
US-09-682-597A-5
; Sequence 5, Application US/09682597A
; Patent No. 6689880
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Chen, Guilan
; APPLICANT: Hironaka, Catherine
; APPLICANT: Zhou, Hua-Ping
; TITLE OF INVENTION: Glyphosate Tolerant Wheat Plant 33391 and Compositions and Method
; FILE REFERENCE: 38-21(52232)A
; CURRENT APPLICATION NUMBER: US/09/682,597A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: m1ec feature
; LOCATION: (1)..(399)
; OTHER INFORMATION: chimeric sequence of wheat genome and transgene insert
US-09-682-597A-5

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e-22; Length 399;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 GCGGCGCGGTTAACAGCTTACTGAGGTCATTCATATGCTTGAGAGAGAGTCCGATA 396
Db 337 GCGGCGCGGTTAACAGCTTACTGAGGTCATTCATATGCTTGAGAGAGAGTCCGATA 396

QY 397 GTC 399
Db 397 GTC 399

RESULT 11
US-09-682-597A-2/c
; Sequence 2, Application US/09682597A
; Patent No. 6689880
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Chen, Guilan
; APPLICANT: Hironaka, Catherine
; APPLICANT: Zhou, Hua-Ping
; TITLE OF INVENTION: Glyphosate Tolerant Wheat Plant 33391 and Compositions and Method
; FILE REFERENCE: 38-21(52232)A
; CURRENT APPLICATION NUMBER: US/09/682,597A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
```

```
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-682-597A-2

Query Match
Best Local Similarity 100.0%; Pred. No. 2.8e-05; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 CATATGCTTGAGAGAGAGTCCGAGATGTC 399
Db 30 CATATGCTTGAGAGAGAGTCCGAGATGTC 1

RESULT 12
US-09-872-051-4/c
; Sequence 4, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32 (nK603) and Composition and Methods for Det
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(29)
; OTHER INFORMATION: fully synthesized
US-09-872-051-4

Query Match
Best Local Similarity 100.0%; Pred. No. 9e-05; Length 29;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 GAGTCGGATGATGCCAAATTAACCAAG 414
Db 29 GAGTCGGATGATGCCAAATTAACCAAG 1

RESULT 13
US-09-872-051-13
; Sequence 13, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32 (nK603) and Composition and Methods for Det
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
```

;; PRIOR APPLICATION NUMBER: 60/241,215
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/240,014
;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 13
;; LENGTH: 22
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; NAME/KEY: source
;; LOCATION: (1)..(22)
;; OTHER INFORMATION: fully synthesized
US-09-872-051-13

Query Match 4.4%; Score 22; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCGATCCAAATCGGACTG 22
DB 1 AATCGATCCAAATCGGACTG 22

RESULT 14
US-09-872-051-14/c
;; Sequence 14, Application US/09872051
;; Patent No. 6825400
;; GENERAL INFORMATION:
;; APPLICANT: Monsanto Co
;; APPLICANT: Behr, Carl
;; APPLICANT: Hironaka, Catherine
;; APPLICANT: Heck, Gregory
;; APPLICANT: You, Jinsong
;; TITLE OF INVENTION: Corn Event PV-ZMG132 (mk603) and Composition and Methods for Detect
;; TITLE OF INVENTION: Thereof
;; FILE REFERENCE: 38-21(52258)B
;; CURRENT APPLICATION NUMBER: US/09/872,051
;; CURRENT FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: 60/213,567
;; PRIOR FILING DATE: 2000-06-22
;; PRIOR APPLICATION NUMBER: 60/241,215
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/240,014
;; PRIOR FILING DATE: 2000-10-13
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 14
;; LENGTH: 22
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; NAME/KEY: source
;; LOCATION: (1)..(22)
;; OTHER INFORMATION: fully synthesized
US-09-872-051-14

Query Match 4.4%; Score 22; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 ATAAAGGTGGCCCAAGTGAA 498
DB 22 ATAAAGGTGGCCCAAGTGAA 1

RESULT 15
US-09-441-340-29
;; Sequence 29, Application US/09441340
;; Patent No. 6448476
;; GENERAL INFORMATION:
;; APPLICANT: Barry, Gerard F.

;; TITLE OF INVENTION: Phosphonate Metabolizing Plants
;; FILE REFERENCE: 38-21(15303)
;; CURRENT APPLICATION NUMBER: US/09/441,340
;; CURRENT FILING DATE: 1999-11-16
;; EARLIER APPLICATION NUMBER: 60/108,763
;; EARLIER FILING DATE: 1998-11-17
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 29
;; LENGTH: 2107
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: expression
;; OTHER INFORMATION: cassette comprising plant operable promoter linked
;; OTHER INFORMATION: to a leader, intron, a sequence encoding an AMPA
;; OTHER INFORMATION: acetyltransferase, and termination sequence
;; FEATURE:
;; NAME/KEY: promoter
;; LOCATION: (26)..(590)
;; FEATURE:
;; NAME/KEY: 5'UTR
;; LOCATION: (615)..(685)
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: (699)..(1148)
;; FEATURE:
;; NAME/KEY: transic peptide
;; LOCATION: (1149)..(1426)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1427)..(1858)
;; FEATURE:
;; NAME/KEY: terminator
;; LOCATION: (1869)..(2102)
US-09-441-340-29

Query Match 4.0%; Score 20; DB 3; Length 2107;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 GCGGCCGCGTTAACACGCTT 356
DB 1 GCGGCCGCGTTAACACGCTT 20

Search completed: February 10, 2005, 08:49:45
Job time : 154.034 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 1222.79 Seconds
(without alignments)
5727.097 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183

Sequence: 1 gacgtatcatgatgagatggg.....aaggaagcgcgagatgacg 1183

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001as.*
5: geneseq2001bs.*
6: geneseq2002as.*
7: geneseq2002bs.*
8: geneseq2003as.*
9: geneseq2003bs.*
10: geneseq2003cs.*
11: geneseq2003ds.*
12: geneseq2004as.*
13: geneseq2004bs.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1183	100.0	1183	ABK15245	Abk15245 Corn nk60
2	202	17.1	2378	ABD01016	Abd01016 Expressio
3	171	14.5	1501	AAc87194	AAc87194 Rice acti
4	171	14.5	1501	AAc89325	AAc89325 Rice acti
5	171	14.5	2181	ADc84543	ADc84543 Rice acti
6	171	14.5	9359	ABK11039	ABK11039 PVDH636 v
7	171	14.5	9359	ABK10687	ABK10687 Transform
8	169	14.3	5118	ADP79330	ADP79330 Plasmid P
9	169	14.3	7545	ACF58343	ACF58343 Nucleotid
10	169	14.3	11643	ACF58344	ACF58344 Nucleotid
11	168	14.2	4032	AAV23237	AAV23237 Plasmid P
12	163	13.8	6865	AAf80287	AAf80287 Nucleotid
13	163	13.8	10003	AAf80295	AAf80295 Nucleotid
14	159	13.4	4774	ADP83768	ADP83768 Benthic
15	147	12.4	349	AAAD51491	AAAD51491 A. tumefa
16	147	12.4	896	ADP83767	ADP83767 Benthic
17	147	12.4	1360	AAAD51480	AAAD51480 Consensus
18	147	12.4	3754	AAZ51642	AAZ51642 Bacillus
19	147	12.4	3754	AAAD61790	AAAD61790 Bt cry3Bb
20	147	12.4	3754	ABX95186	ABX95186 B. thurin

ALIGNMENTS

21	147	12.4	4149	3	AAZ51641	AAZ51641 Bacillus
22	147	12.4	4149	10	AAAD61789	AAAD61789 Bt cry3Bb
23	147	12.4	4149	10	ABX95185	ABX95185 B. thurin
24	147	12.4	8349	3	AAAI5565	AAAI5565 pMON30464
25	147	12.4	10249	3	AAAI5563	AAAI5563 pMON33828
26	147	12.4	10312	3	AAAI5564	AAAI5564 pMON33829
27	147	12.4	10339	3	AAAI5564	AAAI5564 pMON33827
28	144	12.2	6600	12	ADK98482	ADK98482 B thuring
29	144	12.2	7000	12	ADK98485	ADK98485 B thuring
30	140	11.8	2378	3	AAAD1016	AAAD1016 Expressio
31	140	11.8	2436	3	AAAD1018	AAAD1018 Expressio
32	137	11.6	1630	3	AAAD1014	AAAD1014 Expressio
33	137	11.6	6313	6	ABA05258	ABA05258 Human ICA
34	137	11.6	6313	10	ADK97340	ADK97340 Immunoadh
35	137	11.6	8074	6	ABA05262	ABA05262 Plasmid P
36	137	11.6	8074	10	ADK97345	ADK97345 Immunoadh
37	137	11.6	10846	6	ABSS4336	ABSS4336 E. coli f
38	137	11.6	10847	2	AAK08923	AAK08923 Vector co
39	137	11.6	10900	2	AAK08924	AAK08924 Vector co
40	137	11.6	10900	6	ABSS4337	ABSS4337 E. coli f
41	137	11.6	11606	12	ADQ13598	ADQ13598 Plasmid P
42	136	11.5	1691	4	AAAD1910	AAAD1910 pG1 vecto
43	136	11.5	1851	4	AAAD1912	AAAD1912 pG1 vecto
44	136	11.5	2315	4	AAAD1911	AAAD1911 pG1 vecto
45	136	11.5	2708	4	AAAD1908	AAAD1908 pTEN6 vec

RESULT 1
ABK15245
ID ABK15245 standard; DNA, 1183 BP.
XX
AC ABK15245;
XX
DT 08-MAY-2002 (first entry)
XX
DE Corn nk603 3' transgene/genomic sequence PCR fragment.
XX
KW ds; rice actin 1 promoter; RA1; RA1 intron;
KW chloroplast transit peptide gene; glycosylase resistance; corn;
KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;
KW transcripional terminator; cauliflower mosaic virus 35S promoter;
KW PV-ZMG732; transgenic; nk603.
XX
XX Zea mays.
OS Escherichia coli.
OS Agrobacterium tumefaciens.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT misc_feature 1..164
FT FT /*tag= a
FT /*note= "T-AGRTU.nos vector sequence"
FT misc_feature 165..381
FT FT /*tag= b
FT FT /*note= "Vector sequence"
FT misc_feature 382..686
FT FT /*tag= c
FT /*note= "Corn plasmid genes rps11/rpoA partial sequence"
FT FT /*tag= d
FT /*note= "Corn genomic sequence"
XX
XX BP1167531-AL.
XX
XX 02-JAN-2002.
XX
XX 15-JUN-2001; 2001EP-00202314.
XX PF AAZ51642 Bacillus
XX
XX 22-JUN-2000; 2000US-0213567P.

PR 13-OCT-2000; 2000US-0240014P.
 PR 13-OCT-2000; 2000US-0241215P.
 XX
 PA (MONS) MONSANTO TECHNOLOGY LLC.

PI Behr CF, Hironaka C, Heck GR, You J;
 DR WPI; 2002-165871/22.

XX Novel DNA construct useful for producing a corn plant that tolerates
 PT application of glyphosate herbicide, comprises two transgene expression
 PT cassettes.

XX Claim 3; Page 16-17; 25pp; English.

XX The invention relates to a DNA construct (I) comprising a first (F1) and
 CC a second (F2) expression cassette. F1 of the DNA construct, in operable
 CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
 CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
 CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator
 CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,
 CC (V), and (VI). Also included are two DNA sequences from corn plant
 CC PV-ZMT32(hk603) of 498 or 1183bp, or primers and probes derived from
 CC them (used to detect transgene junction points). The construct is useful
 CC for producing a corn plant that tolerates application of glyphosate
 CC herbicide by transforming a corn cell with, selecting the corn cell for
 CC tolerance to application of glyphosate, and growing the corn cell into a
 CC fertile corn plant. The present sequence is the 1183bp amplicon
 CC consisting of the Agrobacterium transcriptional terminator sequence,
 CC vector sequences and corn sequences

XX Sequence 1183 BP; 295 A; 289 G; 280 G; 319 T; 0 U; 0 Other;

Query Match 100.0%; Score 1183; DB 6; Length 1183;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTTATTTATGAGATGGTTTATGATTAGATCCCGCAATTATACATTTAATACGC 60
 DB 1 GAGCTTTATTTATGAGATGGTTTATGATTAGATCCCGCAATTATACATTTAATACGC 60
 QY 61 GATGAAAAAATAATATGCGCGCAACTAGATTAATTTATGCGCGCGGTCTATCTAT 120
 DB 61 GATGAAAAAATAATATGCGCGCAACTAGATTAATTTATGCGCGCGGTCTATCTAT 120
 QY 121 GTTACTAGATCGGGATATCCCGGGGAATTCGGTACCAAGCTTTTATTAATAGTAA 180
 DB 121 GTTACTAGATCGGGATATCCCGGGGAATTCGGTACCAAGCTTTTATTAATAGTAA 180
 QY 181 GAGTAAATTTCACTTTGGGCACTTTTATTAACGATATTTTACTTTATACCACTTTTA 240
 DB 181 GAGTAAATTTCACTTTGGGCACTTTTATTAACGATATTTTACTTTATACCACTTTTA 240
 QY 241 ACTGATGTTTCACTTTGACCAAGTATCTTACTTTGTTTATTTTGGACTATCCGA 300
 DB 241 ACTGATGTTTCACTTTGACCAAGTATCTTACTTTGTTTATTTTGGACTATCCGA 300
 QY 301 CTCTCTCTCAAGATATGATGACCTCGAGTAAGCTTTAAGCGCGCGCCCTAAGGGA 360
 DB 301 CTCTCTCTCAAGATATGATGACCTCGAGTAAGCTTTAAGCGCGCGCCCTAAGGGA 360
 QY 361 TATCAAGCTTGTGATCAACGACCACTTCACTAGTGTGTTGAGTGATCTGTTATC 420
 DB 361 TATCAAGCTTGTGATCAACGACCACTTCACTAGTGTGTTGAGTGATCTGTTATC 420
 QY 421 TCTTCTGAAACATPAACAGACTAGTATTTATGATCATTTAATCTTTCTTTGAA 480
 DB 421 TCTTCTGAAACATPAACAGACTAGTATTTATGATCATTTAATCTTTCTTTGAA 480
 QY 481 AGCGGTTTCAATTTTATTTTACAGACGCTTTTATGAGAGGTGCAATCCATTAATGCGG 540
 DB 481 AGCGGTTTCAATTTTATTTTACAGACGCTTTTATGAGAGGTGCAATCCATTAATGCGG 540

QY 541 ATAGTGTTCATGCGGTATACAACTTAACGGTACACCACTTTAGCAATGCTCGTAA 600
 DB 541 ATAGTGTTCATGCGGTATACAACTTAACGGTACACCACTTTAGCAATGCTCGTAA 600
 QY 601 GCGGATCTCTTCGCTACACGACCTTTTACATATCTTCGCTGTTGCAAAACCACT 660
 DB 601 GCGGATCTCTTCGCTACACGACCTTTTACATATCTTCGCTGTTGCAAAACCACT 660
 QY 661 GTACGAATAGCATCTACTGCTGTTCTGCTACTTTATTTTATTAAGTAAAAACC 720
 DB 661 GTACGAATAGCATCTACTGCTGTTCTGCTACTTTATTTTATTAAGTAAAAACC 720
 QY 721 ATAAATGACACAAACACCTTGCCTTCACTACCTGCGGACGACCGCAAGATGGG 780
 DB 721 ATAAATGACACAAACACCTTGCCTTCACTACCTGCGGACGACCGCAAGATGGG 780
 QY 781 TTCAACACGTCGACACAGATGCAACCGACCTCCAGACCAATCTGAGGCGGAC 840
 DB 781 TTCAACACGTCGACACAGATGCAACCGACCTCCAGACCAATCTGAGGCGGAC 840
 QY 841 GACGACGTAGGACAGGGGTGCGCATTAACGACGTGCGGACATCCACTTGTCTTCC 900
 DB 841 GACGACGTAGGACAGGGGTGCGCATTAACGACGTGCGGACATCCACTTGTCTTCC 900
 QY 901 TCTCTGCTTCACTTCCGCGCGGACGTCTCTAGACCCAGGAGATGCTGTGAGAG 960
 DB 901 TCTCTGCTTCACTTCCGCGCGGACGTCTCTAGACCCAGGAGATGCTGTGAGAG 960
 QY 961 GGTGCGGCGCGCGCATTTTATATGCTTGGCGGACGACCTTGGCGGAACCGATC 1020
 DB 961 GGTGCGGCGCGCGCATTTTATATGCTTGGCGGACGACCTTGGCGGAACCGATC 1020
 QY 1021 GCTCTGCGCAATACAGAAAGACAGTGGGCGCTGCGCTAGGCCACCGCCAGAG 1080
 DB 1021 GCTCTGCGCAATACAGAAAGACAGTGGGCGCTGCGCTAGGCCACCGCCAGAG 1080
 QY 1081 GGGGCTTGTGCGAGCGGTAGCGTGGGAAAGGGACGACCGCTAGGGGGGCCATG 1140
 DB 1081 GGGGCTTGTGCGAGCGGTAGCGTGGGAAAGGGACGACCGCTAGGGGGGCCATG 1140
 QY 1141 CAGGCGCCAGAGAAAAAAGAAAGAGAGCGCGGATGATG 1183
 DB 1141 CAGGCGCCAGAGAAAAAAGAAAGAGAGCGCGGATGATG 1183

RESULT 2
 AAD01016/C
 ID AAD01016 standard; DNA; 2378 BP.
 XX
 AC AAD01016;
 XX
 DT 21-SEP-2000 (first entry)
 XX
 DE Expression cassette-3 comprising modified E. coli p2A phno coding gene.
 XX
 KW Phosphonate herbicide tolerance; aminomethyl phosphonic acid; AMPA; AAT;
 KW acyltransferase; transacylase; recombinant plant; expression cassette;
 KW corn; tobacco; wheat; cotton; canola; rice; chloroplast transit peptide;
 KW CTP; glyphosate oxidase; GOX; glyphosate oxidoreductase; phno gene;
 KW self-fertilisation; hetero-fertilisation; ds.
 XX
 OS Escherichia coli.
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT promoter 28..965
 FT /tag= a
 FT /note= "Plant functional heterologous promoter"
 FT intron 966..1423
 FT /tag= b
 FT /note= "Enhances expression of the phno gene"
 FT transit_peptide 1440..1667
 FT /tag= c

/note="Targets the protein to an intracellular
organelle, like chloroplast"

CDS

1668..2102

/*tag= d

/product= "Escherichia coli modified p2A AMPA

acetyltransferase (AAT) enzyme"
/function= "Transfers acyl group from an acylcarrier
(CoA) to the free amino group of aminomethylphosphonate"
2114..2369

/*tag= e

terminator

MO200029596-A1.

25-MAY-2000.

16-NOV-1999; 99MO-US027152.

17-NOV-1998; 98US-0108763P.

(MONS) MONSANTO CO.

Barry GF;

WPI; 2000-367806/33.

P-PSDB; AAT71251.

Enhancing phosphonate herbicide tolerance in corn, tobacco, wheat,
cotton, canola and rice plants involves transforming plants with
phosphonate metabolizing genes encoding acyltransferase enzyme.

Disclosure; Page 167-169; 179pp; English.

The patent discloses a method for selectively enhancing the phosphonate
herbicide tolerance in recombinant corn, tobacco, wheat, cotton, canola
and rice plants, by transforming the plants with an expression cassette.
It comprises of a structural DNA sequence, that encodes an aminomethyl
phosphonic acid (AMPA) acyltransferase or transacylase (AAT), capable of
N-acylation of AMPA. This sequence has an amino terminal chloroplast
transit peptide (CTP), that targets AAT to the chloroplast. Co-expression
of glyphosate oxidase (GOX) gene, encoding glyphosate oxidoreductase,
along with AAT provides the transformed plants with higher resistance to
phosphonate herbicides. This method is useful for enhancing phosphonate
herbicides such as glyphosate and AMPA tolerance, in recombinant plants
and to prevent self-fertilisation and enhance hetero-fertilisation. The
present DNA sequence is the expression cassette-3, comprising a plant
operable promoter and 5' sequences like intron, linked to modified E.
coli pho coding region, encoding p2A AMPA acyltransferase (AAT) enzyme
and plant operable termination sequences

Sequence 2378 BP; 604 A; 625 C; 552 G; 597 T; 0 U; 0 Other;

Query Match 17.1%; Score 202; DB 3; Length 2378;

Beet Local Similarity 100.0%; Pred. No. 8 7e-90;

Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

163 TTTTATATAGTAGAAGAAAGTAATTTCACTTTGGGCACTTTTATACGATATTTT 222

202 TTTTATATAGTAGAAGAAAGTAATTTCACTTTGGGCACTTTTATACGATATTTT 143

223 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGAACAGATTAATCTTACCTTTGTTT 282

142 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGAACAGATTAATCTTACCTTTGTTT 83

283 TATTTGACATATCCGCACTCTCTCTCAAGCATATGATGACCTCGAGTAGCTTGTGTTA 342

82 TATTTGACATATCCGCACTCTCTCTCAAGCATATGATGACCTCGAGTAGCTTGTGTTA 23

343 ACGCGCGCGCCCTAGGATATC 364

22 ACGCGCGCGCCCTAGGATATC 1

AAc87194/c
ID AAc87194 standard; DNA; 1501 BP.

AAc87194;

09-MAR-2001 (first entry)

Rice actin promoter enhancer element, SEQ ID NO:49.

Rice EPSPS; 5-enolpyruvylshikimate phosphate synthase;
glyphosate resistance; herbicide resistance; transgenic plant;
expression construct; enhancer element; db.

Oryza sativa.

MO200066748-A1.

09-NOV-2000.

20-APR-2000; 2000MO-GB001573.

29-APR-1999; 99GB-00009968.

29-JUL-1999; 99GB-00017834.

29-JUL-1999; 99GB-00017839.

29-JUL-1999; 99GB-00017840.

29-JUL-1999; 99GB-00017846.

29-JUL-1999; 99GB-00017847.

21-DEC-1999; 99GB-00030200.

21-DEC-1999; 99GB-00030204.

21-DEC-1999; 99GB-00030207.

21-DEC-1999; 99GB-00030209.

21-DEC-1999; 99GB-00030213.

(ZENE) ZENECA LTD.

Hawkes TR, Warner SAJ, Andrews CJ, Bachoo S, Pickerill AP;

WPI; 2000-687544/67.

Novel polynucleotide encoding 5-enolpyruvylshikimate phosphate synthase,
PT used to produce transgenic plants e.g. banana, wheat, maize or rice,
having resistance or tolerance to glyphosate herbicide.

Claim 19; Page 55-56; 87pp; English.

The invention relates to rice 5-enolpyruvylshikimate phosphate synthase

(EPSPS) genomic DNA (AAc87188). The invention also relates to an

expression cassette comprising, in the 5'-3' direction, one or more

transcriptional enhancer elements selected from AAc87190-C87196), the

rice EPSPS promoter, genomic DNA encoding a rice EPSPS chloroplast

transit peptide, genomic DNA encoding a EPSPS protein modified such that

it is resistant to glyphosate (AAc87189), and a transcriptional

terminator. The glyphosate resistant EPSPS contains a region (AAB29793)

containing two amino acid substitutions relative to the corresponding

wild-type region (AAB29792). The invention also encompasses plant genomic

EPSPS sequences identified via screening with a rice EPSPS intronic

sequence; vectors and host plant cells comprising a nucleic acid sequence

of the invention; transgenic plants (and tissues and seeds thereof)

comprising a nucleic acid sequence of the invention, optionally further

transformed with a DNA encoding an insect, fungal, viral, bacterial,

nematode, stress or herbicide resistance protein; and methods of

producing the transgenic plants of the invention. The nucleic acids and

constructs of the invention are used to produce a wide variety of

morphologically normal, glyphosate resistant plants. The glyphosate

resistant plants produced are particularly maize, soybean, cotton,

sugarcane and canola, but also other field crops, fruits and vegetables,

turf and forage grasses and nut-producing plants. The plants are

optionally resistant to insects, fungi, viruses, bacteria, nematodes,

stress, desiccation and/or other herbicides. They can be used in the

production of a herbicidal target for the high throughput in vitro

screening of potential herbicides. The present sequence represents an

enhancer element which may be used in the rice EPSPS expression cassette

of the invention

XX Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;
SQ Best Local Similarity 100.0%; Pred. No. 2,4e-74;
Query Match 14.5%; Score 171; DB 3; Length 1501;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 163 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTTATACCGATATTTT 222
DB 920 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTTATACCGATATTTT 861
QY 223 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCAAGTAATCTTACCTTTGTTT 282
DB 860 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCAAGTAATCTTACCTTTGTTT 801
QY 283 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 333
DB 800 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 750

RESULT 4

AAC89325/C
ID AAC89325 standard; DNA; 1501 BP.

AC AAC89325;

DT 07-MAR-2001 (first entry)

DE Rice actin enhancer.

XX 5-enolpyruvylshikimate phosphate synthase; EPSPS; herbicide resistance;
XX glyphosate; ds.

OS Oryza sp.

XX WO200066747-A1.

XX 09-NOV-2000.

XX 20-APR-2000; 2000WO-GB001572.

XX 29-APR-1999; 99GB-00009967.

XX 29-APR-1999; 99GB-00009969.

XX 29-APR-1999; 99GB-00009972.

XX 29-APR-1999; 99GB-00009981.

XX 29-JUL-1999; 99GB-00017835.

XX 29-JUL-1999; 99GB-00017836.

XX 29-JUL-1999; 99GB-00017843.

XX 21-DEC-1999; 99GB-00030202.

XX 21-DEC-1999; 99GB-00030210.

XX 21-DEC-1999; 99GB-00030212.

XX (ZENE) ZENECA LTD.

XX Hawkes TR, Warner SAJ, Andrews CJ, Bachoo S, Pickerill AP;

XX WPI; 2000-679764/66.

XX Isolated polynucleotide encoding a 5-enolpyruvylshikimate phosphate

XX synthase from rice is used for producing transgenic plants with enhanced

XX resistance to glyphosate herbicide.

XX Example 20; Page 54; 98pp; English.

XX The present invention relates to an Oryza sp. 5-enolpyruvylshikimate

XX phosphate synthase (EPSPS) gene. Vectors containing the gene may be used

XX to produce plant tissues and fertile whole plants which are substantially

XX tolerant or substantially resistant to glyphosate herbicide and to

XX produce a herbicidal target which is used for high throughput in vitro

XX screening of potential herbicides

XX Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;

Query Match 14.5%; Score 171; DB 3; Length 1501;
SQ Best Local Similarity 100.0%; Pred. No. 2,4e-74;
Query Match 14.5%; Score 171; DB 3; Length 1501;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 163 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTTATACCGATATTTT 222
DB 920 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTTATACCGATATTTT 861
QY 223 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCAAGTAATCTTACCTTTGTTT 282
DB 860 ACTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCAAGTAATCTTACCTTTGTTT 801
QY 283 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 333
DB 800 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 750

RESULT 5

ADC84543/C
ID ADC84543 standard; DNA; 2181 BP.

AC ADC84543;

DT 01-JAN-2004 (first entry)

DE Rice actin promoter encoding sequence.

XX expression cassette; ds.

XX Melanomy caliginosus.

XX WO2003027257-A2.

XX 03-APR-2003.

XX 27-SEP-2002; 2002WO-US030895.

XX 27-SEP-2001; 2001US-0325607P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Levin JZ, Dietrich R, Budziszewski GJ;

XX WPI; 2003-354651/33.

XX New isolated nucleic acid molecule encoding a polypeptide with a 3'-5'

XX exonuclease domain, useful in molecular biology and transformation, in

XX particular for reproducing and predictably manipulating gene expression

XX in a plant cell.

XX Disclosure; Page 106-107; 108pp; English.

XX The present invention relates to an invention that results in the

XX expression of the endogenous nucleotide sequence in plant cell being

XX increased. In this method the plant cell comprises a second expression

XX cassette comprising a nucleic acid molecule, where the expression of the

XX nucleic acid molecule of interest in the plant cell is decreased as

XX compared to expression of the nucleic acid molecule of interest in a

XX plant cell lacking the first expression cassette. The methods and

XX compositions of the present invention are useful in molecular biology and

XX transformation, in particular for reproducing and predictably

XX manipulating gene expression in a plant cell. The present sequence is a

XX nucleotide sequence of rice actin promoter.

XX Sequence 2181 BP; 602 A; 518 G; 430 C; 631 T; 0 U; 0 Other;

XX Query Match 14.5%; Score 171; DB 10; Length 2181;

XX Best Local Similarity 100.0%; Pred. No. 2,3e-74;

XX Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTTATACCGATATTTT 222

DB 925 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTTATACCGATATTTT 866

CC modification may result in a heritable change in one or more plant
CC characteristics such as inhibition of flowering (or substantial delay
CC that amounts to inhibition), absence of inflorescence, increased
CC production of tillers, delayed heading and inhibition of the
CC developmental switch from vegetative to generative growth. A method of
CC making a grass involves transforming the grass with a nucleic acid which
CC interferes with metabolism of gibberellic acid. A grass can be treated by
CC applying a phytohormone to at least partially relieve or reverse a change
CC in plant characteristic resulting from genetic modification. The
CC genetically modified grass is useful for growing and/or propagating grass
CC in athletic fields (for sports such as baseball, cricket, football, golf,
CC rugby, soccer and tennis), lawns, parks and other types of landscaping.
CC The grass is also useful as an animal feedstuff for cattle, goats, horses
CC and sheep, due to its increased vegetative growth, improved digestibility
CC and/or nutritional value as animal feedstuff. This sequence represents
CC plasmid pVD636 DNA. This plasmid was used as a transformation vector for
CC production of transgenic grasses expressing the Arabidopsis thaliana
CC homeobox gene, AtH1
CC
SQ Sequence 9359 BP; 2377 A; 2222 C; 2133 G; 2624 T; 0 U; 3 Other;

Query Match 14.5%; Score 171; DB 6; Length 9359;
Best Local Similarity 100.0%; Pred. No. 2.2e-74; Indels 0; Gaps 0;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAAGAGTAATTTCACTTTGGCCACCTTTTATACCGATATTT 222
DB 6858 TTTTATATAGTAGAGAAAAGAGTAATTTCACTTTGGCCACCTTTTATACCGATATTT 6799
QY 223 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGAACGAGTAATCTTACCTTTGTT 282
DB 6798 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGAACGAGTAATCTTACCTTTGTT 6739
QY 283 TATTTTGACATATCCGACCTCTCTTCAAGCATATGATGACCTCGAGTA 333
DB 6738 TATTTTGACATATCCGACCTCTCTTCAAGCATATGATGACCTCGAGTA 6688

RESULT 8
ADP73930/C
ID ADP73930 standard; DNA; 5118 BP.
XX
AC ADP73930;
XX
DT 09-SEP-2004 (first entry)
XX
DE Plasmid pDAB3014, SEQ ID NO:84.
XX
KW Transgenic plant; immunoglobulin production; recombinant production;
KW glycosylation; fucose; glycan; virucide; immunotherapy; maize; lipase;
KW 3'UTR; rice; actin promoter; phosphinothricin acyltransferase; PAT;
KW plasmid; pDAB3014; cyclic; circular; ds.
XX
OS Zea mays.
OS Oryza sativa.
OS Chimeric.
OS Synthetic.
OS Unidentified.
XX
FH Key
FH promoter
FT 1172..1724
FT /tag= a
FT /note= "Rice actin promoter"
FT 1727..2281
FT /tag= b
FT /product= "Phosphinothricin acyltransferase (PAT)"
FT 2296..2627
FT /*tag= c
FT /*note= "Maize lipase 3'UTR. The 3' end of this 3'UTR has
FT been deduced from that given for pDAB505 (SEQ ID NO:85),
FT as the 3'UTR location given for this plasmid is 2236-
FT 6652"
XX

PN WO2004050838-A2.
XX
XX 17-JUN-2004.
XX
XX 28-NOV-2003; 2003MO-US037905.
XX
XX 27-NOV-2002; 2002US-0429385P.
XX
XX (DOMC) DOM CHEM CO.
XX (DOMC) DOM AGROSCIENCES LLC.
XX (EPIC-) EPICYTE PHARM INC.
XX
XX Briggs K, Glancy T, Hein MB, Hiatt AC, Karnoup AL, Anderson WHK,
XX Paredy D, Petolino J, Rubin-Wilson B, Taylor D, Roberts JL;
XX WPI, 2004-46111/43.
XX
XX Novel plant-produced immunoglobulin having glycopeptide or glycan profile
XX with reduced fucosylation, useful for treating herpes simplex virus
XX infection.
XX
XX Claim 69; SEQ ID NO 84; 212pp; English.
XX
XX The invention relates to the production of immunoglobulins in plants,
XX wherein at least a portion of the glycans attached to the immunoglobulins
XX lack fucose. The immunoglobulins produced can be of any class (i.e., IgG,
XX IgA, IgM, IgE or IgD) and is especially an anti-herpes simplex virus
XX (HSV) antibody or an anti-alphavetera3, alphavetera5 dual integrin
XX antibody. The invention also relates to constructs, plasmids and vectors
XX for producing the immunoglobulins; transformed plant cells; calli, plant
XX tissues and whole plants for producing the immunoglobulins; methods for
XX producing the immunoglobulins, the immunoglobulins thus produced; and the
XX use of such immunoglobulins. The immunoglobulins of the invention may be
XX used to treat HSV infection or tumour angiogenesis. The invention
XX provides the advantages of antibody production in plants, such as large
XX scale production, reduced costs, and elimination of pathogenic
XX contaminants such as viruses and prions, with a simplified (i.e., non-
XX plant-specific) glycosylation profile which reduces the risk that the
XX immunoglobulin may not be functional in animals. The present sequence
XX represents the plasmid pDAB3014, which may be used in the invention. The
XX plasmid contains a phosphinothricin acyltransferase (PAT) gene under the
XX control of a rice actin promoter, and a maize lipase 3'UTR.
XX
SQ Sequence 5118 BP; 1240 A; 1256 C; 1286 G; 1336 T; 0 U; 0 Other;

Query Match 14.3%; Score 169; DB 12; Length 5118;
Best Local Similarity 100.0%; Pred. No. 2.2e-73; Indels 0; Gaps 0;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAAGAGTAATTTCACTTTGGCCACCTTTTATACCGATATTT 222
DB 496 TTTTATATAGTAGAGAAAAGAGTAATTTCACTTTGGCCACCTTTTATACCGATATTT 437
QY 223 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGAACGAGTAATCTTACCTTTGTT 282
DB 436 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGAACGAGTAATCTTACCTTTGTT 377
QY 283 TATTTTGACATATCCGACCTCTCTTCAAGCATATGATGACCTCGAG 331
DB 376 TATTTTGACATATCCGACCTCTCTTCAAGCATATGATGACCTCGAG 328

RESULT 9
ACF58343/C
ID ACF58343 standard; DNA; 7545 BP.
XX
XX ACF58343;
XX
XX 12-FEB-2004 (first entry)
XX
XX Nucleotide sequence of plasmid pDAB504.
XX
XX GntII, N-acetylglucosaminyltransferase; transgenic; glycoprotein; ds.

```

XX OS Synthetic.
XX FH Key
XX FT misc_feature
XX FT Location/Qualifiers
XX FT 1. .1146
XX FT /tag= a
XX FT /note= "tobacco Rb7 MARS"
XX FT misc_feature
XX FT 1167. .1304
XX FT /tag= b
XX FT /note= "linker sequence (ACF58351)"
XX FT promoter
XX FT 1305. .2701
XX FT /tag= c
XX FT /note= "rice actin promoter and intron"
XX FT intron
XX FT 2235. .2696
XX FT /tag= d
XX FT /note= "rice actin intron"
XX FT misc_feature
XX FT 2704. .3258
XX FT /tag= e
XX FT /note= "PAT gene"
XX FT misc_feature
XX FT 3259. .3272
XX FT /tag= f
XX FT /note= "linker sequence (ACF58352)"
XX FT 3'UTR
XX FT 3273. .3629
XX FT /tag= g
XX FT /note= "maize lipase UTR"
XX FT misc_feature
XX FT complement (3671. .4836)
XX FT /tag= i
XX FT /note= "tobacco Rb7 MAR"
XX FT misc_feature
XX FT 4351. .4405
XX FT /tag= h
XX FT /note= "linker sequence (ACF58353)"
XX FT misc_feature
XX FT 4837. .4857
XX FT /tag= j
XX FT /note= "linker sequence (ACF58354)"
XX FT misc_feature
XX FT 4858. .5103
XX FT /tag= k
XX FT /note= "Puc19"
XX FT misc_feature
XX FT 5104. .5130
XX FT /tag= l
XX FT /note= "linker sequence (ACF58355)"
XX FT misc_feature
XX FT 5130. .7523
XX FT /tag= m
XX FT /note= "Puc19"
XX FT misc_feature
XX FT 7524. .7545
XX FT /tag= n
XX FT /note= "linker sequence (ACF58356)"
XX PN W02003078614-A2.
XX PD 25-SEP-2003.
XX PF 18-MAR-2003; 2003WO-1B001562.
XX PR 19-MAR-2002; 2002US-0365769P.
XX PR 26-MAR-2002; 2002US-0368047P.
XX PA (PLAN-) PLANT RES INT BV.
XX PI Bakker HAC, Florack DEA, Bosch HJ;
XX DR WPI; 2003-779132/73.
XX PT New plant host cell system for producing a desired glycoprotein comprises
XX PT a mammalian N-acetylglucosaminyltransferase (GntIII) enzyme, a nucleic
XX PT acid sequence encoding the enzyme, or a vector comprising the nucleic
XX PT acid sequence.
XX PS Example 6; Fig 7B; 122p; English.
XX CC UDP-N-acetylglucosamine: beta-D-mannoside beta(1,4)-N-
XX CC acetylglucosaminyltransferase (GntIII) enzyme, a nucleic acid sequence
XX CC encoding a mammalian GntIII protein, or a vector comprising the GntIII

```

```

CC nucleic acid. The plant host system is useful in producing a desired
CC glycoprotein or its functional fragment. The glycoprotein or its
CC functional fragment may be used for the production of a pharmaceutical
CC composition. The present sequence represents the nucleotide sequence of
CC plasmid pDA8504
XX SO Sequence 7545 BP; 2128 A; 1579 C; 1622 G; 2216 T; 0 U; 0 Other;
XX
XX Query Match 14.3%; Score 169; DB 10; Length 7545;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-73;
XX Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 163 TTTTAAATAGTAGAAAGAGTAATTTCACTTTGGGCCACTTTTATTAACCATATTTT 222
XX |||||||
XX Db 1473 TTTTAAATAGTAGAAAGAGTAATTTCACTTTGGGCCACTTTTATTAACCATATTTT 1414
XX |||||||
XX QY 223 ACTTTATACCACTTTTAACTGATGTTTCACTTTGACCAAGTAATCTTACCTTTGTT 282
XX |||||||
XX Db 1413 ACTTTATACCACTTTTAACTGATGTTTCACTTTGACCAAGTAATCTTACCTTTGTT 1354
XX |||||||
XX QY 283 TATTTGACTATCCGACTCTCTTCAAGCATATGATGACTCGAG 331
XX |||||||
XX Db 1353 TATTTGACTATCCGACTCTCTTCAAGCATATGATGACTCGAG 1305
XX |||||||
XX
XX RESULT 10
XX ACF58344/c
XX ID ACF58344 standard; DNA; 11643 BP.
XX AC ACF58344;
XX AC ACF58344;
XX DT 12-FEB-2004 (first entry)
XX XX
XX DE Nucleotide sequence of plasmid pDA87113.
XX KM GntIII; N-acetylglucosaminyltransferase; transgenic; glycoprotein; ds.
XX OS Synthetic.
XX FH Key
XX FH Location/Qualifiers
XX FT 1. .1164
XX FT /tag= a
XX FT /note= "Rb7 MAR v3"
XX FT misc_feature
XX FT 1165. .1233
XX FT /tag= b
XX FT /note= "linker sequence (ACF58357)"
XX FT promoter
XX FT 1234. .3224
XX FT /tag= c
XX FT /note= "maize ubiquitin 1 promoter"
XX FT misc_feature
XX FT 3225. .4891
XX FT /tag= d
XX FT /note= "GNTIII v.2"
XX FT 3'UTR
XX FT 4896. .5260
XX FT /tag= e
XX FT /note= "maize peroxidase-5 3'-UTR"
XX FT misc_feature
XX FT 5261. .5404
XX FT /tag= f
XX FT /note= "multiple cloning sites (ACF58358)"
XX FT promoter
XX FT 5405. .6802
XX FT /tag= g
XX FT /note= "rice actin 1 promoter v2"
XX FT misc_feature
XX FT 6803. .7358
XX FT /tag= h
XX FT /note= "PAT v3"
XX FT misc_feature
XX FT 7359. .7372
XX FT /tag= i
XX FT /note= "linker sequence (ACF58359)"
XX FT 3'UTR
XX FT 7373. .7729
XX FT /tag= j
XX FT /note= "maize lipase UTR"
XX FT misc_feature
XX FT 7730. .7770
XX FT /tag= k
XX FT /note= "linker sequence (ACF58360)"

```

```
FT misc_feature 7771..8934
FT /+tag= 1
FT /note= "Rp7 MAR v3"
FT misc_feature 8935..11643
FT /+tag= m
FT /note= "Puc19"
FT misc_feature 10164..11021
FT /+tag= n
FT /note= "ampicillin resistance gene"
XX PN MO2003078614-A2.
XX PD 25-SEP-2003.
XX PF 18-MAR-2003; 2003WO-IB001562.
XX PR 19-MAR-2002; 2002US-0365769P.
XX PR 26-MAR-2002; 2002US-0368047P.
XX PA (PLAN-) PLANT RES INT BV.
XX PI Bakker HAC, Florack DEA, Bosch HJ;
XX DR WPI; 2003-779132/73.
XX PT New plant host cell system for producing a desired glycoprotein comprises
XX PT a mammalian N-acetylglucosaminyltransferase (GntII) enzyme, a nucleic
XX PT acid sequence encoding the enzyme, or a vector comprising the nucleic
XX PT acid sequence.
XX PS Example 6; Fig 8B; 122pp; English.
XX CC The invention relates to a plant host cell system comprising a mammalian
XX CC UDP-N-acetylglucosamine: beta-D-mannoside beta(1,4)-N-
XX CC acetylglucosaminyltransferase (GntII) enzyme, a nucleic acid sequence
XX CC encoding a mammalian GntII protein, or a vector comprising the GntII
XX CC nucleic acid. The plant host system is useful in producing a desired
XX CC glycoprotein or its functional fragment. The glycoprotein or its
XX CC functional fragment may be used for the production of a pharmaceutical
XX CC composition. The present sequence represents the nucleotide sequence of
XX CC plasmid pDA87113
XX SQ Sequence 11643 BP; 3005 A; 2672 C; 2635 G; 3331 T; 0 U; 0 Other;
XX
XX Query Match 14.3%; Score 169; DB 10; Length 11643;
XX Best Local Similarity 100.0%; Pred. No. 2,1e-73;
XX Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 163 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCCACCTTTATTACCGATATTTT 222
XX DB 5573 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCCACCTTTATTACCGATATTTT 5514
XX
XX QY 223 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCGAGTAATTTTACCTTTGTTT 282
XX DB 5513 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCGAGTAATTTTACCTTTGTTT 5454
XX
XX QY 283 TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGATGACCTCGAG 331
XX DB 5453 TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGATGACCTCGAG 5405
XX
XX RESULT 11
XX AAV23237/c
XX ID AAV23237 standard; DNA; 4032 BP.
XX AC AAV23237;
XX AC 17-JUL-1998 (first entry)
XX DE Plasmid pMV71.
XX XX Barstar; barnase inhibitor; fertility restoration; male-sterile line;
XX KW plasmid pMV71; circular; ds.
```

```
XX OS Synthetic.
XX FH Key
XX FT promoter
XX FT CDS
XX FT 3'UTR
XX PN WO9810081-A2.
XX PD 12-MAR-1998.
XX PF 01-SEP-1997; 97WO-EP004739.
XX PF 03-SEP-1996; 96EP-00202446.
XX PR (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX PI Michiels F, Williams M;
XX DR WPI; 1998-193630/17.
XX PT DNA encoding an improved barstar protein - used to restore fertility in
XX PT male-sterile plant lines.
XX PS Example 3; Page 37-39; 54pp; English.
XX CC The present sequence was used in the preparation of an improved Bacillus
XX CC amyloliquefaciens barstar, i.e. barnase inhibitor, which can be used to
XX CC restore fertility to male-sterile lines. The DNA sequence encoding the
XX CC improved barstar, leads to increased barstar production in tapetum cells,
XX CC due to improved translation, and possibly protein stability
XX CC
XX SQ Sequence 4032 BP; 1072 A; 968 C; 963 G; 1029 T; 0 U; 0 Other;
XX
XX Query Match 14.2%; Score 168; DB 2; Length 4032;
XX Best Local Similarity 100.0%; Pred. No. 7e-73;
XX Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 163 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCCACCTTTATTACCGATATTTT 222
XX DB 2166 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCCACCTTTATTACCGATATTTT 2107
XX
XX QY 223 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCGAGTAATTTTACCTTTGTTT 282
XX DB 2106 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTTGACCGAGTAATTTTACCTTTGTTT 2047
XX
XX QY 283 TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGATGACCTCGA 330
XX DB 2046 TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGATGACCTCGA 1999
XX
XX RESULT 12
XX AAF80287
XX ID AAF80287 standard; DNA; 6865 BP.
XX AC AAF80287;
XX AC 29-JUN-2001 (first entry)
XX DE Nucleotide sequence of plasmid pMR1195.
XX XX Vector; transgenesis; trfA locus; RK2 ori; oriV; p285 protein;
XX KW P382 protein; antibiotic resistance gene; nptII; transgenic plant; ss.
```

OS	Synthetic.
XX	
FH	Key
FT	rep_origin
FT	
FT	rep_origin
FT	
CDS	
FT	
FT	
FT	CDS
FT	
FT	
FT	misc_feature
FT	
FT	terminator
FT	
FT	CDS
FT	
FT	intron
FT	
FT	promoter
FT	
FT	misc_feature
FT	
PN	FR2798139-A1.
PD	09-MAR-2001.
XX	
Pf	03-SEP-1999; 99PR-00011112.
XX	
PR	03-SEP-1999; 99PR-00011112.
XX	(MERI-) MERISTEM THERAPEUTICS SA.
PA	
PI	Gruber V, Comeau D;
DR	WPI; 2001-259847/27.
XX	
PT	New vector free from non-essential elements, useful for transforming
PT	cells for protein production and for preparing transgenic plants.
PS	Claim 20; Page 128-131; 180pp; French.
XX	The specification describes a synthetic vector containing only those
CC	elements essential for its functionality and transgenesis of a cell
CC	(especially a plant cell). The vector consists of at most one origin of
CC	replication (ori), at most one sequence encoding a selection agent and a
CC	tirA locus encoding a protein that increases the level of plasmid
CC	replication. The vector particularly contains an RK2 ori, especially orally
CC	from PRK2 of Escherichia coli with a broad host range, an antibiotic
CC	resistance gene (especially nptII conferring resistance to kanamycin in
CC	bacterial) and a tirA locus from PRK2 encoding the proteins P285 and P382
CC	The vectors are used to prepare transgenic plants and transformed host
CC	cells for production of a heterologous proteins, e.g. insulin,
CC	interferon, lipase, blood proteins and anti-inflammatory agents. The
CC	present sequence represents a plasmid of the invention
XX	
Sequence	6865 BP; 1699 A; 1694 C; 1919 G; 1553 T; 0 U; 0 Other;

Query Match 13.8%; Score 163; DB 4; Length 6865;

	Best Local Similarity	100.0%;	Pred. No. 2.1e-70;	Matches 163;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	163	TTTTATATATGTAAGAAAGTAATTTCACTTTGGGCGACCTTTTATTAACGATATTTT	222					
Db	6383	TTTTATATATAGTAGAAAGAGTAATTTCACTTTGGGCGACCTTTTATTAACGATATTTT	6442					
OY	223	ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAAGTAATCTTACCTTTGTTT	282					
Db	6443	ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAAGTAATCTTACCTTTGTTT	6502					
OY	283	TATTTTGACATATCCGACCTCTCTTTCACGACATATGAATGAC	325					
Db	6503	TATTTTGACATATCCGACCTCTCTTTCACGACATATGAATGAC	6545					
RESULT 13								
AF80295	ID	AAF80295 standard; DNA; 10003 BP.						
XX	AC	AAF80295;						
XX	DT	29-JUN-2001 (first entry)						
XX	DE	Nucleotide sequence of plasmid pNR11210.						
XX	XX	Vector; transgenesis; trfA locus; RK2 ori; oriV; P285 protein;						
KW	P382	protein; antibiotic resistance gene; nptIII; transgenic plant; ss.						
XX	OS	Synthetic.						
XX	Key	Location/Qualifiers						
FH	FT	1..654						
FH	FT	/*tag= a						
FT	FT	/note= "ori RK2"						
FT	FT	655..1263						
FT	FT	/*tag= b						
FT	FT	/note= "ori ColE1"						
FT	FT	1264..2603						
FT	FT	/*tag= c						
FT	FT	/note= "NPT III gene coding for neomycin phosphotransferase and kanamycin resistance"						
FT	FT	2604..4098						
FT	FT	/*tag= d						
FT	FT	/note= "trfA locus from RK2 coding for two proteins P285 and P382 enabling an increase in the replication rate"						
FT	FT	4106..4271						
FT	FT	/*tag= e						
FT	FT	/note= "T-DNA left border"						
FT	FT	4272..4559						
FT	FT	/*tag= f						
FT	FT	/note= "nopaline synthetase terminator"						
FT	FT	4594..5169						
FT	FT	/*tag= g						
FT	FT	/note= "Bar gene coding for phosphinotricin acetyltransferase and glufosinate resistance"						
FT	FT	5170..5704						
FT	FT	/*tag= h						
FT	FT	/note= "rice actin intron"						
FT	FT	5705..6638						
FT	FT	/*tag= i						
FT	FT	/note= "rice actin promoter"						
FT	FT	6672..7107						
FT	FT	/*tag= j						
FT	FT	/note= "wheat high molecular weight glutenin promoter"						
FT	FT	7169..7687						
FT	FT	/*tag= k						
FT	FT	/note= "rice actin intron"						
FT	FT	7688..9496						
FT	FT	/*tag= l						
FT	FT	/note= "GUS gene coding for beta glucuronidase"						
FT	FT	9497..9823						
FT	FT	/*tag= m						

XX Huber SA, Roberts JK, Shapley ZW, Doherty S;
XX WPI; 2003-148719/14.
XX

PT Insect resistant cotton plants, tissues and seeds that include the
MON15985 event, useful in plant insect protection and plant breeding.

XX Disclosure; Page 47; 52pp; English.

XX The invention relates to insect resistant cotton plants, tissues and
CC seeds that include the MON15985 event. The methods and compositions of
CC the invention are useful in the field of plant molecular biology, in
CC particular plant insect protection and plant breeding. The MON15985 event
CC confers resistance to lepidopteran insect damage. The present sequence is
CC A. tumefaciens DNA used in the exemplification of the invention
XX

SQ Sequence 349 BP; 101 A; 68 C; 74 G; 106 T; 0 U; 0 Other;

Query Match 12.4%; Score 147; DB 8; Length 349;

Best Local Similarity 100.0%; Pred. No. 2.2e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGTGGGTTTTATGATTAGAGTCCCGCAATTATACATTATATAGC 60
DB 179 GACGTTATTATGAGTGGGTTTTATGATTAGAGTCCCGCAATTATACATTATATAGC 238
QY 61 GATGAAACAAATATAGCGCGCAACTAGATTAATATCGCGCGGTGCATCTAT 120
DB 239 GATGAAACAAATATAGCGCGCAACTAGATTAATATCGCGCGGTGCATCTAT 298
QY 121 GTTACTAGATCGGGGATATCCCGGGG 147
DB 299 GTTACTAGATCGGGGATATCCCGGGG 325

Search completed: February 10, 2005, 13:44:01
Job time : 1225.79 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 1967.4 Seconds
(without alignments)
3460.943 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183
Sequence: 1 gacgtatctatcgagatggg.....aaggaagcgcgagatgatgc 1183

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4313806 segs, 2877871033 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1183	100.0	1183	9 US-09-872-051-8	Sequence 8, Appli
2	1183	100.0	1183	18 US-10-790-430-8	Sequence 8, Appli
3	305	25.8	6683	18 US-10-425-115-58081	Sequence 58081, A
4	202	17.1	2378	14 US-10-213-791-27	Sequence 27, Appl
5	171	14.5	1501	14 US-10-012-070A-49	Sequence 49, Appl
6	171	14.5	1501	14 US-10-012-013-44	Sequence 44, Appl
7	169	14.3	2480	19 US-10-678-588A-1	Sequence 1, Appli
8	168	14.2	4032	9 US-09-970-921-5	Sequence 5, Appli
9	163	13.8	6865	10 US-09-845-064-13	Sequence 13, Appl
10	163	13.8	10003	10 US-09-845-064-21	Sequence 21, Appl
11	147	12.4	349	18 US-10-480-223A-23	Sequence 23, Appl

12	147	12.4	1360	18 US-10-480-223A-12	Sequence 12, Appl
13	147	12.4	3754	15 US-10-233-665-15	Sequence 15, Appl
14	147	12.4	4149	15 US-10-233-665-13	Sequence 13, Appl
15	147	12.4	8349	16 US-10-198-478-16	Sequence 16, Appl
16	147	12.4	10249	16 US-10-198-478-14	Sequence 15, Appl
17	147	12.4	10312	16 US-10-198-478-15	Sequence 15, Appl
18	147	12.4	10339	16 US-10-198-478-13	Sequence 13, Appl
19	144	12.2	11546	18 US-10-841-796-33	Sequence 27, Appl
20	140	11.8	2378	15 US-10-213-791-27	Sequence 31, Appl
21	140	11.8	2436	15 US-10-213-791-31	Sequence 23, Appl
22	137	11.6	1630	15 US-10-213-791-23	Sequence 9, Appli
23	137	11.6	6313	13 US-10-047-542-9	Sequence 14, Appl
24	137	11.6	8074	13 US-10-047-542-14	Sequence 5, Appli
25	137	11.6	10846	9 US-09-923-109-5	Sequence 5, Appli
26	137	11.6	10846	15 US-10-164-204-5	Sequence 5, Appli
27	137	11.6	10846	17 US-10-705-430-5	Sequence 6, Appli
28	137	11.6	10900	9 US-09-923-109-6	Sequence 6, Appli
29	137	11.6	10900	15 US-10-164-204-6	Sequence 6, Appli
30	137	11.6	10900	17 US-10-705-430-6	Sequence 6, Appli
31	137	11.6	11606	18 US-10-602-475A-9	Sequence 69, Appl
32	136	11.5	1691	18 US-10-168-653C-69	Sequence 71, Appl
33	136	11.5	1851	18 US-10-168-653C-71	Sequence 70, Appl
34	136	11.5	2315	18 US-10-168-653C-70	Sequence 67, Appl
35	136	11.5	2708	18 US-10-168-653C-67	Sequence 19, Appl
36	136	11.5	11522	14 US-10-052-092-19	Sequence 19, Appl
37	136	11.5	11522	16 US-10-437-107-19	Sequence 19, Appl
38	136	11.5	11522	18 US-10-896-419-19	Sequence 104, App
39	134	11.3	277	15 US-10-161-403-104	Sequence 107, App
40	134	11.3	277	18 US-10-161-403-107	Sequence 15, Appl
41	134	11.3	277	18 US-10-161-408-15	Sequence 18, Appl
42	134	11.3	277	18 US-10-161-408-18	Sequence 6, Appli
43	134	11.3	563	9 US-09-970-921-6	Sequence 16, Appl
44	134	11.3	1213	9 US-09-794-384A-16	Sequence 16, Appl
45	134	11.3	1213	18 US-10-801-550-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-872-051-8
Sequence 8, Application US/09872051
Patent No. US20020013960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinong
TITLE OF INVENTION: Corn Event PV-ZMGT32(Ink603) and Composition and Methods for D
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1.164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea maize plastid gene, rps11 and rpoA
OTHER INFORMATION: 687-1183 Zea maize genomic DNA

US-09-872-051-8

Query Match	100.0%;	Score 1183;	DB 9;	Length 1183;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1183; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Oy	1	GACGTTATTATGACATGGGTTTTTATGATATGAGTCCCGCAATATACATTTAATAGC	60
Db	1	GACGTTATTATGACATGGGTTTTTATGATATGAGTCCCGCAATATACATTTAATAGC	60
Oy	61	GATGAAAAACAAATATATAGCGGCACAACCTAGATTAATATTCGCGGGGTGCATCTAT	120
Db	61	GATGAAAAACAAATATATAGCGGCACAACCTAGATTAATATTCGCGGGGTGCATCTAT	120
Oy	121	GTACTAATATCGGGGATATCCCCGGGGGAATTCGGTACCAAGTTTATATATATGATGAAA	180
Db	121	GTACTAATATCGGGGATATCCCCGGGGGAATTCGGTACCAAGTTTATATATATGATGAAA	180
Oy	181	GAGTAAATTTCACTTTGGGCGCACCTTTTATATACGATATTTTATCTTTATACACCTTTTA	240
Db	181	GAGTAAATTTCACTTTGGGCGCACCTTTTATATACGATATTTTATCTTTATACACCTTTTA	240
Oy	241	ACTGATGTTTTCACTTTTGAACGAGTAATCTTACTTTGTTTATATTTTGAAGTATCCGA	300
Db	241	ACTGATGTTTTCACTTTTGAACGAGTAATCTTACTTTGTTTATATTTTGAAGTATCCGA	300
Oy	301	CTCTCTTCTCAAGCATATGAATAGACCTTGAGTAACCTTGTAACGCGGCGCCCTTAGGGA	360
Db	301	CTCTCTTCTCAAGCATATGAATAGACCTTGAGTAACCTTGTAACGCGGCGCCCTTAGGGA	360
Oy	361	TATCAAGCTTGGTATCACGCGCACACACTTCCACTAGTGTGTTGAGTGAATCTGTATTC	420
Db	361	TATCAAGCTTGGTATCACGCGCACACACTTCCACTAGTGTGTTGAGTGAATCTGTATTC	420
Oy	421	TCTTCTCGAACCATACAGACTAGTATATTTGATCATTTGAATGTTTATTTCTCTGAA	480
Db	421	TCTTCTCGAACCATACAGACTAGTATATTTGATCATTTGAATGTTTATTTCTCTGAA	480
Oy	481	AGCGGTTTCATTTTTTTTATAGAGAGTCTTTTTTATAGAGGTCGACATCCATATAGGGG	540
Db	481	AGCGGTTTCATTTTTTTTATAGAGAGTCTTTTTTATAGAGGTCGACATCCATATAGGGG	540
Oy	541	ATAGGTGTTACATCGCGTATACAACTTAAACCGTACACCACTTTTATAGCAATGGCTGTAT	600
Db	541	ATAGGTGTTACATCGCGTATACAACTTAAACCGTACACCACTTTTATAGCAATGGCTGTAT	600
Oy	601	GCGGATCTCTTCGCGTACACAGACCTTTTACATTAACCTTGCTGCTGTTCMAAACCACT	660
Db	601	GCGGATCTCTTCGCGTACACAGACCTTTTACATTAACCTTGCTGCTGTTCMAAACCACT	660
Oy	661	GTACGATATGACATCTACTGCTGTCTGCTGACTTATATTTTTTTTATATATGABAAAAAC	720
Db	661	GTACGATATGACATCTACTGCTGTCTGCTGACTTATATATTTTTTTTATATATGABAAAAAC	720
Oy	721	ATAAATATGACAACAACACCTGCCCTTCACTACCGGTCGAGCGACGCGGAAGATGGGG	780
Db	721	ATAAATATGACAACAACACCTGCCCTTCACTACCGGTCGAGCGACGCGGAAGATGGGG	780
Oy	781	TTTCAACAGGCTGCGCACACAGATGCAACGGAACCTTCCAAAGCCAAATATCTCGAGGCGGAC	840
Db	781	TTTCAACAGGCTGCGCACACAGATGCAACGGAACCTTCCAAAGCCAAATATCTCGAGGCGGAC	840
Oy	841	GACGACGTATAGGAGGGGTGGCCATATAGCAGCGGTGGCGGATTCACACTTGTTCTTCCCTT	900
Db	841	GACGACGTATAGGAGGGGTGGCCATATAGCAGCGGTGGCGGATTCACACTTGTTCTTCCCTT	900
Oy	901	TCTCTGTCTTCAACTTTCGCGCGGCGAGTCTGTAGACCAAGGGGATGCTGTGTGAGAGAGA	960
Db	901	TCTCTGTCTTCAACTTTCGCGCGGCGAGTCTGTAGACCAAGGGGATGCTGTGTGAGAGAGA	960
Oy	961	GATCGCGGGGCCCCATTTTTTATATGCTGTGGGCGAGACGACTTGGCGGAACCGATCCAGA	1020
Db	961	GATCGCGGGGCCCCATTTTTTATATGCTGTGGGCGAGACGACTTGGCGGAACCGATCCAGA	1020

Qy	1021	GCTCTGGGCAAAATACAGAAAGAACGATGGGGGCGCTGCGCGCTAAGGCCACCGCAGGAGAC	1080
Db	1021	GCTCTGGGCAAAATACAGAAAGAACGATGGGGGCGCTGCGCGCTAAGGCCACCGCAGGAGAC	1080
Qy	1081	GGGGCTTGTTGCGAGCGCGTAGCGTGGGAAAGGGGACGACC CGCTAAGGGGGGCCATGCTC	1140
Db	1081	GGGGCTTGTTGCGAGCGCGTAGCGTGGGAAAGGGGACGACC CGCTAAGGGGGGCCATGCTC	1140
Qy	1081	GGGGCTTGTTGCGAGCGCGTAGCGTGGGAAAGGGGACGACC CGCTAAGGGGGGCCATGCTC	1140
Db	1081	GGGGCTTGTTGCGAGCGCGTAGCGTGGGAAAGGGGACGACC CGCTAAGGGGGGCCATGCTC	1140
Qy	1141	CAGCGCCAGAGAAAAAAGAAAGGAGCGCGAGATGATG	1183
Db	1141	CAGCGCCAGAGAAAAAAGAAAGGAGCGCGAGATGATG	1183

RESULT 2
US-10-790-430-8

Sequence 8, Application US/10790430
Publication No. US20040139493A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMGR32(nk603) and Composition and Methods for Detection and Identification of the Same
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/10/790,430
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea mays plastid gene, rps11 and rpoA
OTHER INFORMATION: 687-1183 Zea mays genomic DNA
US-10-790-430-8

[illegible]

```
Db 241 ACTGATGTTTCACTTTGACGAGTATCTTACTTTGTTATTTGACATACCGA 300
Qy 301 CTCTCTTCACCATATGATGACCTCGAGTAAAGCTTTGTAACGGCGCCCTAGGA 360
Db 301 CTCTCTTCACCATATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy 361 TATCAAGCTTGTACCAACGCGGACACCTTCAGTATGTTGATGATGATGATGAT 420
Db 361 TATCAAGCTTGTACCAACGCGGACACCTTCAGTATGTTGATGATGATGATGAT 420
Qy 421 TCTTCTCGAACCAATACAGACTAGTATGATGATGATGATGATGATGATGATGAT 480
Db 421 TCTTCTCGAACCAATACAGACTAGTATGATGATGATGATGATGATGATGATGAT 480
Qy 481 AGCGGTTTCACTTTTTCACAGACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAG 540
Db 481 AGCGGTTTCACTTTTTCACAGACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAG 540
Qy 541 ATAGGTGTTACATCGCTATACAACTTAACCGTACACCACTTTTTCAGAGCTTT 600
Db 541 ATAGGTGTTACATCGCTATACAACTTAACCGTACACCACTTTTTCAGAGCTTT 600
Qy 601 GCGGACCTCTTCGCTACCAAGCACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAG 660
Db 601 GCGGACCTCTTCGCTACCAAGCACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAG 660
Qy 661 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 721 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 781 TTCAACAGCTGCGGACCAAGCACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAG 840
Db 781 TTCAACAGCTGCGGACCAAGCACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAG 840
Qy 841 GAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 GAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 901 TCTCTGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 901 TCTCTGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Qy 961 GGTGCGGCGGCGCGGATTTTTCAGAGCTTTTTCAGAGCTTTTTCAGAGCTTT 1020
Db 961 GGTGCGGCGGCGCGGATTTTTCAGAGCTTTTTCAGAGCTTTTTCAGAGCTTT 1020
Qy 1021 GCTCTGCGCAATACAGAGCACTGATGATGATGATGATGATGATGATGATGAT 1080
Db 1021 GCTCTGCGCAATACAGAGCACTGATGATGATGATGATGATGATGATGATGAT 1080
Qy 1081 GGGGCTTGTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 GGGGCTTGTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy 1141 CAGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
Db 1141 CAGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
```

RESULT 3

```
US-10-425-115-58081/c
; Sequence 58081, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 58081
; LENGTH: 6683
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRL4577_152966C.1
US-10-425-115-58081
Query Match 25.8%; Score 305; DB 18; Length 6683;
Best Local Similarity 100.0%; Pred. No. 3,1e-149; Indels 0; Gaps 0;
Matches 305; Conservative 0; Mismatches 0;
Qy 382 ACACTTCCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 441
Db 5673 ACACTTCCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5614
Qy 442 TAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
Db 5613 TAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5554
Qy 502 AGAGCTTTTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 561
Db 5553 AGAGCTTTTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 5494
Qy 562 CAACCTTACCGTACCACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAGAGCTTT 621
Db 5493 CAACCTTACCGTACCACTTTTTCAGAGCTTTTTCAGAGCTTTTTCAGAGCTTT 5434
Qy 622 GCACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
Db 5433 GCACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5374
Qy 682 GTTCT 686
Db 5373 GTTCT 5369
```

RESULT 4

```
US-10-213-791-27/c
; Sequence 27, Application US/10213791
; Publication No. US20030106096A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15103)
; CURRENT APPLICATION NUMBER: US/10/213,791
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US/09/441,340
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/108,763
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression
; OTHER INFORMATION: cassette comprising a plant promoter linked to an
; OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
; OTHER INFORMATION: transferase, and a termination sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (28)..(965)
; FEATURE:
; NAME/KEY: intron
```

LOCATION: (966) .. (1423)
FEATURE:
NAME/KEY: transit_peptide
LOCATION: (1440) .. (1667)
FEATURE:
NAME/KEY: CDS
LOCATION: (1668) .. (2099)
FEATURE:
NAME/KEY: terminator
LOCATION: (2114) .. (2369)
US-10-213-791-27

Query Match 17.1%; Score 202; DB 15; Length 2378;
Best Local Similarity 100.0%; Pred. No. 4.2e-95;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAGATAATTTCATTGGGCGACCTTTTATACCATATTTT 222
DB 202 TTTTATATAGTAGAGAAAGATAATTTCATTGGGCGACCTTTTATACCATATTTT 143
QY 223 ACTTATACCACTTTTAACTGATGTTTTCATTGGGCGACCTTTTATACCATATTTT 282
DB 142 ACTTATACCACTTTTAACTGATGTTTTCATTGGGCGACCTTTTATACCATATTTT 83
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTT 342
DB 82 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTT 23
QY 343 ACGCGCGCGCCCTTAGGATATC 364
DB 22 ACGCGCGCGCCCTTAGGATATC 1

RESULT 5
US-10-012-070A-49/C
Sequence 49, Application US/10012070A
Publication No. US20030077801A1
GENERAL INFORMATION:
APPLICANT: Hawkes, Timothy
APPLICANT: Warner, Simon
APPLICANT: Andrews, Christopher
APPLICANT: Bachoo, Satvinder
APPLICANT: Pickerill, Andrew
TITLE OF INVENTION: Herbicide Resistant Plants
FILE REFERENCE: 50490/US
CURRENT APPLICATION NUMBER: US/10/012,070A
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: PCT/GB00/01573
PRIOR FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 49
LENGTH: 1501
TYPE: DNA
ORGANISM: Oryza sp.
US-10-012-070A-49

Query Match 14.5%; Score 171; DB 14; Length 1501;
Best Local Similarity 100.0%; Pred. No. 8.2e-79;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAGATAATTTCATTGGGCGACCTTTTATACCATATTTT 222
DB 920 TTTTATATAGTAGAGAAAGATAATTTCATTGGGCGACCTTTTATACCATATTTT 861
QY 223 ACTTATACCACTTTTAACTGATGTTTTCATTGGGCGACCTTTTATACCATATTTT 282
DB 860 ACTTATACCACTTTTAACTGATGTTTTCATTGGGCGACCTTTTATACCATATTTT 801
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTT 333
DB 800 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTT 750

RESULT 6
US-10-012-013-44/C
Sequence 44, Application US/10012013
Publication No. US20030079246A1
GENERAL INFORMATION:
APPLICANT: Hawkes, Timothy
APPLICANT: Warner, Simon
APPLICANT: Andrews, Christopher
APPLICANT: Bachoo, Satvinder
APPLICANT: Pickerill, Andrew
TITLE OF INVENTION: Herbicide Resistant Plants
FILE REFERENCE: 50450/US
CURRENT APPLICATION NUMBER: US/10/012,013
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: PCT/GB00/01572
PRIOR FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 44
LENGTH: 1501
TYPE: DNA
ORGANISM: Oryza sp.
US-10-012-013-44

Query Match 14.5%; Score 171; DB 14; Length 1501;
Best Local Similarity 100.0%; Pred. No. 8.2e-79;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAGATAATTTCATTGGGCGACCTTTTATACCATATTTT 222
DB 920 TTTTATATAGTAGAGAAAGATAATTTCATTGGGCGACCTTTTATACCATATTTT 861
QY 223 ACTTATACCACTTTTAACTGATGTTTTCATTGGGCGACCTTTTATACCATATTTT 282
DB 860 ACTTATACCACTTTTAACTGATGTTTTCATTGGGCGACCTTTTATACCATATTTT 801
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTT 333
DB 800 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTT 750

RESULT 7
US-10-678-588A-1
Sequence 1, Application US/10678588A
Publication No. US20050022266A1
GENERAL INFORMATION:
APPLICANT: Wu, Jingrui
TITLE OF INVENTION: Water-Deficit-Tolerant Transgenic Plants
FILE REFERENCE: 38-21(52578)C
CURRENT APPLICATION NUMBER: US/10/678,588A
CURRENT FILING DATE: 2003-10-02
PRIOR APPLICATION NUMBER: US 60/415,758
PRIOR FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: US 60/425,157
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: US 60/463,787
PRIOR FILING DATE: 2003-04-11
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 2480
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: transcriptional unit comprising promoter, coding sequence for
OTHER INFORMATION: transcription factor of SEQ ID NO:2 and terminator elements
US-10-678-588A-1

Query Match 14.3%; Score 169; DB 19; Length 2480;
Best Local Similarity 100.0%; Pred. No. 9.2e-78;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      163  TTTTATAATAGTGAAGAGTAATTCACTTTGGGCGACCTTTATTACGATATTTT 222
Db      2312  TTTTATAATAGTGAAGAGTAATTCACTTTGGGCGACCTTTATTACGATATTT 2371
Oy      223  ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAGGTAATCTTACCTTTGTT 282
Db      2372  ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAGGTAATCTTACCTTTGTT 2431
Oy      283  TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAG 331
Db      2432  TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAG 2480

RESULT 8
US-09-970-921-5/c
; Sequence 5, Application US/09970921
; Patent No. US20020134845A1
GENERAL INFORMATION:
APPLICANT: Frank Michiels et al.
TITLE OF INVENTION: Improved Barstar Gene
FILE REFERENCE: 2428-0108P
CURRENT APPLICATION NUMBER: US/09/970, 921
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4032
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: other nucleic
OTHER INFORMATION: acid, "plasmid pmv71"
NAME/KEY: misc_feature
LOCATION: (1995)..(3400)
OTHER INFORMATION: label = PRAC1, "promoter region of rice actin gene"
NAME/KEY: misc_feature
LOCATION: (3401)..(3676)
OTHER INFORMATION: label = barstar, "barstar DNA"
NAME/KEY: misc_feature
LOCATION: (3677)..(4003)
OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium"
NAME/KEY: misc_feature
LOCATION: (3395)..(3404)
OTHER INFORMATION: label = NcoI, "NcoI recognition site"
NAME/KEY: misc_feature
LOCATION: (4015)..(4021)
OTHER INFORMATION: label = KpnI, "KpnI recognition site"
US-09-970-921-5

Query Match      14.2%: Score 168; DB 9; Length 4032;
Best Local Similarity 100.0%; Pred. No. 3.1e-77; Indels 0; Gaps 0;
Matches 168; Conservative 0; Mismatches 0;

Oy      163  TTTTATAATAGTGAAGAGTAATTCACTTTGGGCGACCTTTATTACGATATTTT 222
Db      2166  TTTTATAATAGTGAAGAGTAATTCACTTTGGGCGACCTTTATTACGATATTTT 2107
Oy      223  ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAGGTAATCTTACCTTTGTT 282
Db      2106  ACTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAGGTAATCTTACCTTTGTT 2047
Oy      283  TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAG 330
Db      2046  TATTTTGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAG 1999

RESULT 9
US-09-845-064-13
; Sequence 13, Application US/09845064
; Publication No. US20030175976A1

```

```

1  GENERAL INFORMATION: THERAPEUTICS
2  APPLICANT: MERISEM
3  TITLE OF INVENTION: CLEAN SYNTHETIC VECTORS, PLASMIDS, TRANSGENIC PLANTS
4  TITLE OF INVENTION: AND PLANT PARTS CONTAINING SAID VECTORS, AND THEIR
5  TITLE OF INVENTION: METHODS OF PRODUCTION
6  FILE REFERENCE: SynVec1
7  CURRENT APPLICATION NUMBER: US/09/845,064
8  CURRENT FILING DATE: 2001-04-27
9  NUMBER OF SEQ ID NOS: 57
10 SOFTWARE: PatentIn Ver. 2.1
11
12 SEQ ID NO 13
13 LENGTH: 6865
14 TYPE: DNA
15 ORGANISM: Artificial Sequence
16 FEATURE: Description of Artificial Sequence:Plasmid
17 OTHER INFORMATION: pmr1195
18 FEATURE:
19 NAME/KEY: rep_origin
20 LOCATION: (1) - (654)
21 OTHER INFORMATION: Origin of replication ori RK2
22 FEATURE:
23 NAME/KEY: rep_origin
24 LOCATION: (655) ..(1263)
25 OTHER INFORMATION: Origin of replication ori COL1
26 FEATURE:
27 NAME/KEY: gene
28 LOCATION: (1264) .. (2603)
29 OTHER INFORMATION: NPT III gene coding for neomycin
30 OTHER INFORMATION: phosphotransferase and kanamycin resistance
31 FEATURE:
32 NAME/KEY: misc feature
33 LOCATION: (2604) ..(4098)
34 OTHER INFORMATION: TrfA locus from RK2 coding for two proteins, p285
35 OTHER INFORMATION: and p382, enabling the increase in the replication
36 OTHER INFORMATION: rate
37 FEATURE:
38 NAME/KEY: misc feature
39 LOCATION: (4106) ..(4271)
40 OTHER INFORMATION: T-DNA left border
41 FEATURE:
42 NAME/KEY: terminator
43 LOCATION: (4272) ..(4559)
44 OTHER INFORMATION: No. US20030175976A1:alaine synthetase terminator
45 FEATURE:
46 NAME/KEY: gene
47 LOCATION: (4575) ..(5150)
48 OTHER INFORMATION: Bar gene coding for phosphinotricin
49 OTHER INFORMATION: acetyltransferase and glufosinate resistance
50 FEATURE:
51 NAME/KEY: intron
52 LOCATION: (5174) ..(5685)
53 OTHER INFORMATION: Rice Actin Intron
54 FEATURE:
55 NAME/KEY: promoter
56 LOCATION: (5686) ..(6626)
57 OTHER INFORMATION: Rice Actin promoter
58 FEATURE:
59 NAME/KEY: misc feature
60 LOCATION: (6626) ..(6685)
61 OTHER INFORMATION: MCS multiple cloning site
62 FEATURE:
63 NAME/KEY: misc_feature
64 LOCATION: (6685) ..(6858)
65 OTHER INFORMATION: T-DNA right border
66
67 US-09-845-064-13
68
69 Query Match 13.8%; Score 163; DB 10; Length 6865;
70 Best Local Similarity 100.0%; Pred. No. 1.3e-74;
71 Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0
72
73 163 TTTATATATGTGTACAAAAGAGTAATTTGACTTGGGGCCACCTTTATTTACGATATTTT 222
74 |||||

```

```

Db      6383  TTTTATATAGTAAAGAAAGATTAATTCACTTTGGGCGACCTTTATTAACGATATTTT 6442
Oy      223  ACTTATATCCACCTTTTAACTGATGTTTCACTTTTGACGAGTATCTTACCTTGT 282
Db      6443  ACTTATATCCACCTTTTAACTGATGTTTCACTTTTGACGAGTATCTTACCTTGT 6502
Oy      283  TATTTTGACTATCCCGACTCTCTTCTCAAGCATATGATGAC 325
Db      6503  TATTTTGACTATCCCGACTCTCTTCTCAAGCATATGATGAC 6545

RESULT 10
US-09-845-064-21
; Sequence 21, Application US/09845064
; Publication No. US20030175976A1
; GENERAL INFORMATION:
; APPLICANT: MERISTEM THERAPEUTICS
; TITLE OF INVENTION: CLEAN SYNTHETIC VECTORS, PLASMIDS, TRANSGENIC PLANTS
; TITLE OF INVENTION: AND PLANT PARTS CONTAINING SAID VECTORS, AND THEIR
; TITLE OF INVENTION: METHODS OF PRODUCTION
; FILE REFERENCE: Synvec1
; CURRENT APPLICATION NUMBER: US/09/845,064
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 10003
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
; OTHER INFORMATION: pMR1210
; FEATURE:
; NAME/KEY: rep_origin
; LOCATION: (1)..(654)
; OTHER INFORMATION: Origin of replication ori RK2
; FEATURE:
; NAME/KEY: rep_origin
; LOCATION: (655)..(1263)
; OTHER INFORMATION: Origin of replication ori ColEI
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1264)..(2603)
; OTHER INFORMATION: NPT III gene coding for neomycin
; OTHER INFORMATION: phosphotransferase and kanamycin resistance
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2604)..(4098)
; OTHER INFORMATION: TrfA locus from RK2 coding for two proteins, P285
; OTHER INFORMATION: and P382, enabling the increase of the replication
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4106)..(4271)
; OTHER INFORMATION: T-DNA left border
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (4272)..(4559)
; OTHER INFORMATION: No. US20030175976A1AlaIine synthetase terminator
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4559)..(4572)
; OTHER INFORMATION: MCS multiple cloning site
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4594)..(5169)
; OTHER INFORMATION: Bar gene coding for phosphinothricine
; OTHER INFORMATION: acetyltransferase and glutosinate resistance
; FEATURE:
; NAME/KEY: intron
; LOCATION: (5170)..(5704)
; OTHER INFORMATION: Rice Actin Intron
; FEATURE:
;

```

```

NAME/KEY: promoter
LOCATION: (5705)..(6638)
OTHER INFORMATION: Rice Actin Promoter
FEATURE:
NAME/KEY: misc_feature
LOCATION: (6646)..(6672)
OTHER INFORMATION: MCS multiple cloning site
FEATURE:
NAME/KEY: promoter
LOCATION: (6672)..(7107)
OTHER INFORMATION: High Molecular Weight Glutenin promoter from wheat
FEATURE:
NAME/KEY: misc_feature
LOCATION: (7107)..(7169)
OTHER INFORMATION: MCS multiple cloning site
FEATURE:
NAME/KEY: intron
LOCATION: (7169)..(7687)
OTHER INFORMATION: Rice Actin Intron
FEATURE:
NAME/KEY: gene
LOCATION: (7688)..(9496)
OTHER INFORMATION: GUS gene coding for beta glucuronidase
FEATURE:
NAME/KEY: terminator
LOCATION: (9497)..(9823)
OTHER INFORMATION: No. US20030175976A1aline synthetase terminator
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9823)..(9996)
OTHER INFORMATION: T-DNA right border
CS-09-845-064-21

```

```

Query Match 163; Score 163; DB 10; Length 10003;
Best Local Similarity 100.0%; Pred. No. 1,36-74;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 163 TTTATATATAGTAAAGATAATTTCACTTTGGGACCTTTATTTACCGATATTT 222
Db 6402 TTTTATATATATAGAAAGATAATTTCACTTTGGGACCTTTATTTACCGATATTT 6461

Qy 223 ACTTATATACACCTTTTACTGATGTTTTCACCTTTGACACGATATCTTACCTTTGTTT 282
Db 6462 ACTTATATACACCTTTTACTGATGTTTTCACCTTTGACACGATATCTTACCTTTGTTT 6521

Qy 283 TATTTTGACTATCCCGACTCTCTTCTCAAGCATATGATATGC 325
Db 6522 TATTTTGACTATCCCGACTCTCTTCTCAAGCATATGATATGC 6564

RESULT 11
US-10-480-223A-23
; Sequence 23, Application US/10480223A
; Publication No. US20040250317A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto, Technology LLC
; APPLICANT: Hubert, Scott A
; APPLICANT: Doherty, Sean C.
; APPLICANT: Roberts, James K.
; APPLICANT: Shapley, Zachary W.
; TITLE OF INVENTION: Cotton Event 15985 and Compositions and Methods for Detection
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 11899, 0232, PCUS00
; CURRENT APPLICATION NUMBER: US/10/480, 223A
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US 60/297406
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PCT/US 02/17853
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 349

```

```

; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-480-223A-23

Query Match      12.4%; Score 147; DB 18; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.3e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
    |||
Db 179 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 238

Qy 61 GATGAAAAAATAATATAGCGCGCAAACTAGAGATTAATATATCGCGCGGTGCATCTAT 120
    |||
Db 239 GATGAAAAAATAATATAGCGCGCAAACTAGAGATTAATATATCGCGCGGTGCATCTAT 298

Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
    |||
Db 299 GTTACTAGATCGGGGATATCCCGGGG 325

RESULT 12
US-10-480-223A-12
; Sequence 12, Application US/10480223A
; Publication No. US20040250317A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto, Technology LLC
; APPLICANT: Huber, Scott A
; APPLICANT: Doherty, Sean C.
; APPLICANT: Roberts, James K.
; APPLICANT: Shapley, Zachary W.
; TITLE OF INVENTION: Cotton Event 15985 and Compositions and Methods for Detection
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 11899.0232.PCUS00
; CURRENT APPLICATION NUMBER: US/10/480.223A.
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US 60/297406
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PCT/US 02/17853
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(349)
; OTHER INFORMATION: 3' end insert sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (350)..(673)
; OTHER INFORMATION: 3' end Gossypium hirsutum remnant DNA sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674)..(1360)
; OTHER INFORMATION: 3' end Gossypium hirsutum chromosomal sequence
US-10-480-223A-12

Query Match      12.4%; Score 147; DB 18; Length 1360;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
    |||
Db 179 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 238

Qy 61 GATGAAAAAATAATATAGCGCGCAAACTAGAGATTAATATATCGCGCGGTGCATCTAT 120
    |||
Db 239 GATGAAAAAATAATATAGCGCGCAAACTAGAGATTAATATATCGCGCGGTGCATCTAT 298

Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
```

```

Db 299 GTTACTAGATCGGGGATATCCCGGGG 325

RESULT 13
US-10-232-665-15
; Sequence 15, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232.665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377.466
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 3754
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression
; OTHER INFORMATION: cassette
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (25)..(640)
; OTHER INFORMATION: P-CamV.35S
; FEATURE:
; NAME/KEY: intron
; LOCATION: (669)..(1472)
; OTHER INFORMATION: I-Zm.Hsp70
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1490)..(3448)
; OTHER INFORMATION: Cry3Bb1 variant v11231
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (3475)..(3730)
; OTHER INFORMATION: Agrobacterium tumefaciens nos 3' transcription
; OTHER INFORMATION: termination and polyadenylation sequence
US-10-232-665-15

Query Match      12.4%; Score 147; DB 15; Length 3754;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
    |||
Db 3601 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 3660

Qy 61 GATGAAAAAATAATATAGCGCGCAAACTAGAGATTAATATATCGCGCGGTGCATCTAT 120
    |||
Db 3661 GATGAAAAAATAATATAGCGCGCAAACTAGAGATTAATATATCGCGCGGTGCATCTAT 3720

Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
    |||
Db 3721 GTTACTAGATCGGGGATATCCCGGGG 3747

RESULT 14
US-10-232-665-13
; Sequence 13, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232.665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377.466
```

```
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 13
/ LENGTH: 4149
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: expression
/ FEATURE:
/ OTHER INFORMATION: cassette
/ FEATURE:
/ NAME/KEY: promoter
/ LOCATION: (25)..(640)
/ OTHER INFORMATION: P-CanV.35S
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (669)..(1472)
/ OTHER INFORMATION: I-Zm.Hsp70
/ FEATURE:
/ NAME/KEY: transit_peptide
/ LOCATION: (1489)..(1635)
/ OTHER INFORMATION: amino terminal TS-Zm.rbcS
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (1636)..(1798)
/ OTHER INFORMATION: I-Zm.rbcS
/ FEATURE:
/ NAME/KEY: transit_peptide
/ LOCATION: (1799)..(1885)
/ OTHER INFORMATION: carboxy terminus TS-Zm.rbcS
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1885)..(3843)
/ OTHER INFORMATION: Cry3Bb1 variant v11231
/ FEATURE:
/ NAME/KEY: terminator
/ LOCATION: (3871)..(4127)
/ OTHER INFORMATION: T-AGRTu.nos 3' transcription termination and
/ OTHER INFORMATION: polyadenylation sequence
US-10-232-665-13
```

```
Query Match 12.4%; Score 147; DB 15; Length 4149;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 60
DB 3996 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 4055
QY 61 GATGAAAACAAATATAGCGCGCAACTAGATTAATTAATGCGCGGTGTCTAT 120
DB 4056 GATGAAAACAAATATAGCGCGCAACTAGATTAATTAATGCGCGGTGTCTAT 4115
QY 121 GTTACTGATCGGGGATATCCCGGGG 147
DB 4116 GTTACTGATCGGGGATATCCCGGGG 4142
```

RESULT 15

```
US-10-198-478-16
/ Sequence 16, Application US/10198478
/ Publication No. US20030188336A1
/ GENERAL INFORMATION:
/ APPLICANT: Corbin, David R.
/ APPLICANT: Romano, Charles P.
/ TITLE OF INVENTION: Improved Methods for Transforming Plants to Express delta-Endotox
/ FILE REFERENCE: 38-21 (15547) B
/ CURRENT APPLICATION NUMBER: US/10/198,478
/ PRIOR APPLICATION NUMBER: 09/186, 002
/ PRIOR FILING DATE: 1998-11-04
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 16
```

```
/ LENGTH: 8349
/ TYPE: DNA
/ ORGANISM: artificial sequence
/ FEATURE:
/ OTHER INFORMATION: completely synthesized
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(8349)
/ OTHER INFORMATION:
US-10-198-478-16
```

```
Query Match 12.4%; Score 147; DB 16; Length 8349;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 60
DB 5722 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 5781
QY 61 GATGAAAACAAATATAGCGCGCAACTAGATTAATTAATGCGCGGTGTCTAT 120
DB 5782 GATGAAAACAAATATAGCGCGCAACTAGATTAATTAATGCGCGGTGTCTAT 5841
QY 121 GTTACTGATCGGGGATATCCCGGGG 147
DB 5842 GTTACTGATCGGGGATATCCCGGGG 5868
```

```
Search completed: February 9, 2005, 11:12:28
Job time : 1969.4 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 18.6055 Seconds
(without alignments)
5727.097 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18
Sequence: 1 tgcgtctcgtcgtacttc 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0	18	6	ABK15248	Abk15248 Corn tran
2	100.0	183	6	ABK15245	Abk15245 Corn nk60
3	88.9	1361	5	AAS66560	Aas66560 DNA encod
4	88.9	1587	4	AAH14053	Aah14053 Human cDN
5	83.3	21	3	AAZ73249	Aaz73249 Human b1a
6	83.3	445	5	ABV19889	Abv19889 Human pro
7	83.3	500	5	ABV49652	Abv49652 Human pro
8	83.3	715	4	ABL09089	Ab109089 Drosophi1
9	83.3	2280	10	ADES9493	Ades9493 Human gen
10	83.3	2280	10	ADES9491	Ades9491 Human gen
11	83.3	3167	4	AAS30480	Aas30480 DNA encod
12	83.3	3167	4	AAI06260	Aai06260 Human rep
13	83.3	3190	4	ABL09088	Ab109088 Drosophi1
14	83.3	5919	10	ADFS9752	Adfs9752 Human con
15	83.3	41335	4	AAK86202	Aak86202 Human imm
16	83.3	62130	12	ADQ97885	Adq97885 Mouse can
17	83.3	70019	13	ABD33601	Abd33601 Murine ca
18	77.8	20	2	AAK96131	Aak96131 PCR prime
19	77.8	25	12	ADP16460	Adp16460 Renal cel
20	77.8	25	12	ADP16461	Adp16461 Renal cel

ALIGNMENTS

21	14	77.8	27	10	ADD40762	Add40762 Murine te
22	14	77.8	149	5	ABV15475	Abv15475 Human pro
23	14	77.8	154	5	ABV15397	Abv15397 Human pro
24	14	77.8	169	4	ABA47800	Ab47800 Human bre
25	14	77.8	169	4	AAK14094	Aak14094 Human bra
26	14	77.8	169	6	ABS13933	Abs13933 Human gen
27	14	77.8	230	5	ABV36189	Abv36189 Human pro
28	14	77.8	241	10	ACA55805	Aca55805 Zebrafish
29	14	77.8	241	12	AD155601	Ad155601 Human pol
30	14	77.8	300	2	AAZ14788	Aaz14788 Human gen
31	14	77.8	366	8	ACA44008	Aca44008 Prokaryot
32	14	77.8	379	5	ABV36267	Abv36267 Human pro
33	14	77.8	400	9	ACH46396	Ach46396 Human inf
34	14	77.8	426	8	ACA35231	Aca35231 Prokaryot
35	14	77.8	459	11	ACH94209	Ach94209 Klebsiell
36	14	77.8	483	12	ACH90762	Ach90762 Human gen
37	14	77.8	496	4	ABA42671	Ab42671 Human bre
38	14	77.8	496	4	AAK01345	Aak01345 Human bra
39	14	77.8	496	5	ABV45241	Abv45241 Human pro
40	14	77.8	496	5	ABV45307	Abv45307 Human pro
41	14	77.8	496	6	ABS01398	Ab01398 Human ORF
42	14	77.8	515	6	ABN79319	Abn79319 Human cer
43	14	77.8	543	4	AAH70725	Aah70725 Human cer
44	14	77.8	547	12	ACH77045	Ach77045 Human gen
45	14	77.8	549	3	AAC93637	Aac93637 Cat file

RESULT 1

ABK15248	ABK15248 standard; DNA; 18 BP.
AC	ABK15248;
XX	08-MAY-2002 (first entry)
DT	Corn transgene junction probe #3.
XX	Probe; 88; rice actin 1 promoter; RAI; RAI intron;
DE	chloroplast transit peptide gene; glyphosate resistance; corn;
KW	5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;
KW	transcriptional terminator; cauliflower mosaic virus 35S promoter;
KW	PV-ZMGFS2; transgenic; nk603.
XX	Zea mays.
OS	Synthetic.
XX	EP1167531-A1.
PN	02-JAN-2002.
PD	15-JUN-2001; 2001EP-00202314.
XX	22-JUN-2000; 2000US-0213567P.
PR	13-OCT-2000; 2000US-0240014P.
PR	13-OCT-2000; 2000US-0241215P.
XX	(MONS) MONSANTO TECHNOLOGY LLC.
PA	Behr CF, Hironaka C, Heck GR, You J;
XX	WPI; 2002-165871/22.
DR	Novel DNA construct useful for producing a corn plant that tolerates
PT	application of glyphosate herbicide, comprises two transgene expression
PT	cassettes.
FT	Claim 8; Page 3; 25pp; English.
XX	The invention relates to a DNA construct (I) comprising a first (F1) and
XX	a second (F2) expression cassette. F1 of the DNA construct, in operable
CC	

CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,
CC (III), (III) and (IV). Also included are two DNA sequences from corn plant
CC PV-ZMG732(nk603) of 498 or 1183bp, or primers and probes derived from
CC them (used to detect transgene junction points). The construct is useful
CC for producing a corn plant that tolerates application of glyphosate
CC herbicide by transforming a corn cell with, selecting the corn cell into a
CC tolerance to application of glyphosate, and growing the corn cell into a
CC fertile corn plant. The present sequence is a transgene junction probe
CC sequence consisting of corn genomic DNA and vector sequences
CC
SQ Sequence 18 BP; 1 A; 4 C; 4 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 18; Conservative 0; Indels 0; Gaps 0;
OY 1 TGCTGTTCTGCTGACTTT 18
1 TGCTGTTCTGCTGACTTT 18
Db
RESULT 2
ABK15245
ID ABK15245 standard; DNA; 1183 BP.
XX
AC ABK15245;
XX
DT 08-MAY-2002 (first entry)
XX
DE Corn nk603 3' transgene/genomic sequence PCR fragment.
XX
KM ds; rice actin 1 promoter; RA1; RA1 intron;
KM chloroplast transit peptide gene; glyphosate resistance; corn;
KM 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;
KM transcriptional terminator; cauliflower mosaic virus 35S promoter;
KM PV-ZMG732; transgenic; nk603.
XX
OS Zea mays.
OS Escherichia coli.
OS Agrobacterium tumefaciens.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT misc_feature 1..164
FT misc_feature /*tag= a
FT misc_feature /note= "T-AGRTU.nos vector sequence"
FT misc_feature 165..381
FT misc_feature /*tag= b
FT misc_feature /note= "Vector sequence"
FT misc_feature 382..686
FT misc_feature /*tag= c
FT misc_feature /note= "Corn plasmid genes rps11/rpoA partial sequence"
FT misc_feature 687..1183
FT misc_feature /*tag= d
FT misc_feature /note= "Corn genomic sequence"
XX
FN EPI167531-A1.
XX
PD 02-JAN-2002.
XX
PF 15-JUN-2001; 2001EP-00202314.
XX
PR 22-JUN-2000; 2000US-0213567P.
PR 13-OCT-2000; 2000US-0240014P.
PR 13-OCT-2000; 2000US-0241215P.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Behr CF, Hironaka C, Heck GR, You J;

XX
DR WPI; 2002-165871/22.
XX
PT Novel DNA construct useful for producing a corn plant that tolerates
PT application of glyphosate herbicide, comprises two transgene expression
PT cassettes.
XX
PS Claim 3; Page 16-17; 25pp; English.
XX
CC The invention relates to a DNA construct (I) comprising a first (F1) and
CC a second (F2) expression cassette. F1 of the DNA construct, in operable
CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,
CC PV-ZMG732(nk603) of 498 or 1183bp, or primers and probes derived from corn plant
CC them (used to detect transgene junction points). The construct is useful
CC for producing a corn plant that tolerates application of glyphosate
CC herbicide by transforming a corn cell with, selecting the corn cell into a
CC tolerance to application of glyphosate, and growing the corn cell into a
CC fertile corn plant. The present sequence is the 1183bp amplicon
CC consisting of the Agrobacterium transcriptional terminator sequence,
CC vector sequences and corn sequences
CC
SQ Sequence 1183 BP; 295 A; 289 C; 319 T; 0 U; 0 Other;
Query Match 100.0%; Score 18; DB 6; Length 1183;
Best Local Similarity 100.0%; Pred. No. 2.1; Mismatches 0; Gaps 0;
Matches 18; Conservative 0; Indels 0; Gaps 0;
OY 1 TGCTGTTCTGCTGACTTT 18
678 TGCTGTTCTGCTGACTTT 695
Db
RESULT 3
AAS66560
ID AAS66560 standard; cDNA; 1361 BP.
XX
AC AAS66560;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #2364.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG02373.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 2364; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed CC genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC useful for generating antibodies against it, detecting or quantitating a CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic CC coding sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1361 BP; 414 A; 322 C; 312 G; 313 T; 0 U; 0 Other;

XX Query Match 88.9%; Score 16; DB 5; Length 1361;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACT 16
Db 948 TGCTGTTCTGCTGACT 963

RESULT 4
AAH14053
ID AAH14053 standard; cDNA; 1587 BP.
XX AAH14053;
AC 26-JUN-2001 (first entry)
DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:11181.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
OS Homo sapiens.
XX EP1074617-A2.
PN 07-FEB-2001.
PD 28-JUL-2000; 2000EP-00116126.
PF 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
DR WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
PS Claim 8; SEQ ID NO 11181; 2537bp + Sequence listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dt primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AA895893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX SQ Sequence 1587 BP; 312 A; 416 C; 383 G; 476 T; 0 U; 0 Other;

XX Query Match 88.9%; Score 16; DB 4; Length 1587;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACT 16
Db 948 TGCTGTTCTGCTGACT 963

RESULT 5
AAZ73249
ID AAZ73249 standard; DNA; 21 BP.
XX AAZ73249;
AC 10-SEP-2001 (first entry)
DT 10-SEP-2001 (first entry)
XX Human biallelic marker upstream amplification primer SEQ ID NO:7605.
DE Human genome; biallelic marker; high density disequilibrium map;
XX genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation;
KW amplification; single nucleotide polymorphism; SNP; PCR primer;
XX diagnosis; ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO954500-A2.
FN 28-OCT-1999.
XX 21-APR-1999; 99WO-IB000822.
PF 21-APR-1998; 98US-0082614P.
XX 23-NOV-1998; 98US-0109732P.
PR (GEST) GENSET.
PA Cohen D, Blumenfeld M, Chumakov I;
PI WPI; 2000-013267/01.
XX WPI; 2000-013267/01.
DR WPI; 2000-013267/01.
XX Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome.
XX Claim 9; Page 1852; 2745bp; English.

CC AA265654 to AA269578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AA269579 to AA277440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the invention
CC have a variety of uses: they can be used for high density mapping of the
CC human genome, and in complex association studies and haplotyping studies
CC which are useful in determining the genetic basis for disease states.
CC Compositions and methods of the invention can also be useful for the
CC identification of the targets for the development of pharmaceutical
CC agents and diagnostic methods, as well as the characterisation of the
CC differential efficacious responses to and side effects from
CC pharmaceutical agents acting on a disease as well as other treatment.
CC N.B. The SEQ ID Nos 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC 3367, are not actually given a sequence in the Sequence Listing from the
CC present invention
XX
SQ Sequence 21 BP, 1 A; 5 C; 6 G; 9 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTGTTCTGCTGACTT 17
DB 2 CTGTTCTGCTGACTT 16

RESULT 6
ABV19889/c
ID ABV19889 standard; cDNA; 445 BP.
XX
XX ABV19889;
XX
DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 19880.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3245; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 445 BP, 123 A; 97 C; 102 G; 123 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 5; Length 445;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGAC 15
DB 256 TGCTGTTCTGCTGAC 242

RESULT 7
ABV49652/c
ID ABV49652 standard; cDNA; 500 BP.
XX
XX ABV49652;
XX
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 49643.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 9689; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX of progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 500 BP, 139 A, 112 C; 116 G; 133 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 5; Length 500;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TGCTGTTCTGCTGAC 15
Db 296 TGCTGTTCTGCTGAC 282

RESULT 8

ABL09089

ID ABL09089 standard; cDNA; 715 BP.

AC ABL09089;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 21749.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW,

XX WPI; 2001-656860/75.

XX P-PSDB; ABB64986.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 21749; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB160511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AB857737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 715 BP; 151 A; 192 C; 202 G; 170 T; 0 U; 0 Other;

XX Query Match 83.3%; Score 15; DB 4; Length 715;
XX Best Local Similarity 100.0%; Pred. No. 91;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 TGCTGTTCTGCTGAC 15

XX 88 TGCTGTTCTGCTGAC 102

RESULT 9

ADE59493/C

ID ADE59493 standard; DNA; 2280 BP.

XX ADE59493;

XX 29-JAN-2004 (first entry)
XX Human gene AK001865, SEQ ID NO 5388.
XX Human; ds; gene; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AK001865.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or more of the polypeptides given in the
CC activity in an animal of one or more of the polypeptides useful in treating
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2280 BP; 680 A; 358 C; 392 G; 850 T; 0 U; 0 Other;

XX Query Match 83.3%; Score 15; DB 10; Length 2280;
XX Best Local Similarity 100.0%; Pred. No. 97;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 4 TGTTCTGCTGACTTT 18

XX 741 TGTTCTGCTGACTTT 727

RESULT 10
ADES9491/c
ID ADE59491 standard; DNA; 2280 BP.
XX
AC ADE59491;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human gene AK001865, SEQ ID NO 5386.
XX
KW Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003;
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M,
XX
DR GENBANK; AK001865.
XX
PT MPI; 2003-268312/26.
XX
New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
XX
Claim 1; Page; 1017p; English.
XX
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human DNA (shown in Table 2 of the
specification) which encodes one of the polypeptides of the invention
which is differentially expressed during pain. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 4 TGTTCGTGACTTT 18
|||||
Db 741 TGTTCGTGACTTT 727
RESULT 11
AAS30480
ID AAS30480 standard; DNA; 3167 BP.
XX
AC AAS30480;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding novel prostate gland antigen, Seq ID NO 338.
XX
KW Human; nocotropic; neuroprotective; cytostatic; antiapoptotic;
KW antihaemic; dermatological; immunosuppressive; antiinflammatory;
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;
KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;
KW prostatosis; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;
KW reproductive system disorder; autoimmune disorder; urinary system;
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;
KW blood-related disorder; hyperproliferative disorder; respiratory;
KW neurological disorder; endocrine disorder; inflammatory disorder;
KW liver disorder; wound healing; food preservative; ds.
XX
OS Homo sapiens.
XX
PN WO200155447-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001330.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-019076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209457P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0225797P.
PR 18-AUG-2000; 2000US-0226275P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.

PR	-SEP-2000	2000US-0228287P
PR	01-SEP-2000	2000US-0228344B
PR	01-SEP-2000	2000US-0229345P
PR	01-SEP-2000	2000US-0229350P
PR	05-SEP-2000	2000US-0229509P
PR	06-SEP-2000	2000US-0229513P
PR	06-SEP-2000	2000US-0230437P
PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231242P
PR	08-SEP-2000	2000US-0231243P
PR	08-SEP-2000	2000US-0231244P
PR	08-SEP-2000	2000US-0231413P
PR	08-SEP-2000	2000US-0231414P
PR	08-SEP-2000	2000US-0233080P
PR	08-SEP-2000	2000US-0233081P
PR	12-SEP-2000	2000US-0231968P
PR	14-SEP-2000	2000US-0232397P
PR	14-SEP-2000	2000US-0232398P
PR	14-SEP-2000	2000US-0232399P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0233063P
PR	14-SEP-2000	2000US-0233064P
PR	14-SEP-2000	2000US-0233065P
PR	21-SEP-2000	2000US-0234423P
PR	21-SEP-2000	2000US-0234774P
PR	25-SEP-2000	2000US-0234997P
PR	25-SEP-2000	2000US-0234998P
PR	25-SEP-2000	2000US-0234998P
PR	26-SEP-2000	2000US-0235488P
PR	27-SEP-2000	2000US-0235634P
PR	27-SEP-2000	2000US-0235635P
PR	29-SEP-2000	2000US-0236527P
PR	29-SEP-2000	2000US-0236567P
PR	29-SEP-2000	2000US-0236568P
PR	29-SEP-2000	2000US-0236569P
PR	29-SEP-2000	2000US-0236570P
PR	02-OCT-2000	2000US-0236602P
PR	02-OCT-2000	2000US-0237053P
PR	02-OCT-2000	2000US-0237058P
PR	02-OCT-2000	2000US-0237039P
PR	02-OCT-2000	2000US-0237039P
PR	02-OCT-2000	2000US-0237040P
PR	13-OCT-2000	2000US-0239933P
PR	13-OCT-2000	2000US-0239933P
PR	13-OCT-2000	2000US-0239937P
PR	20-OCT-2000	2000US-0241212P
PR	20-OCT-2000	2000US-0241212P
PR	20-OCT-2000	2000US-0241185P
PR	20-OCT-2000	2000US-0241786P
PR	20-OCT-2000	2000US-0241787P
PR	20-OCT-2000	2000US-0241808P
PR	20-OCT-2000	2000US-0241809P
PR	20-OCT-2000	2000US-0241865P
PR	01-NOV-2000	2000US-0244617P
PR	08-NOV-2000	2000US-0246474P
PR	08-NOV-2000	2000US-0246475P
PR	08-NOV-2000	2000US-0246476P
PR	08-NOV-2000	2000US-0246477P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246523P
PR	08-NOV-2000	2000US-0246524P
PR	08-NOV-2000	2000US-0246525P
PR	08-NOV-2000	2000US-0246526P
PR	08-NOV-2000	2000US-0246527P
PR	08-NOV-2000	2000US-0246528P
PR	08-NOV-2000	2000US-0246528P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246609P
PR	08-NOV-2000	2000US-0246610P
PR	08-NOV-2000	2000US-0246611P
PR	17-NOV-2000	2000US-0249207P
PR	17-NOV-2000	2000US-0249208P
PR	17-NOV-2000	2000US-0249209P
PR	17-NOV-2000	2000US-0249210P
PR	17-NOV-2000	2000US-0249211P

XX	17-NOV-2000;	2000US-0249212P.	PR			
XX	17-NOV-2000;	2000US-0249213P.	PR			
XX	17-NOV-2000;	2000US-0249214P.	PR			
XX	17-NOV-2000;	2000US-0249215P.	PR			
XX	17-NOV-2000;	2000US-0249216P.	PR			
XX	17-NOV-2000;	2000US-0249217P.	PR			
XX	17-NOV-2000;	2000US-0249218P.	PR			
XX	17-NOV-2000;	2000US-0249244P.	PR			
XX	17-NOV-2000;	2000US-0249245P.	PR			
XX	17-NOV-2000;	2000US-0249264P.	PR			
XX	17-NOV-2000;	2000US-0249265P.	PR			
XX	17-NOV-2000;	2000US-0249297P.	PR			
XX	17-NOV-2000;	2000US-0249299P.	PR			
XX	17-NOV-2000;	2000US-0249300P.	PR			
XX	01-DEC-2000;	2000US-0250160P.	PR			
XX	01-DEC-2000;	2000US-0250391P.	PR			
XX	05-DEC-2000;	2000US-0251030P.	PR			
XX	05-DEC-2000;	2000US-0251988P.	PR			
XX	05-DEC-2000;	2000US-0256719P.	PR			
XX	06-DEC-2000;	2000US-0251479P.	PR			
XX	08-DEC-2000;	2000US-0251856P.	PR			
XX	08-DEC-2000;	2000US-0251868P.	PR			
XX	08-DEC-2000;	2000US-0251869P.	PR			
XX	08-DEC-2000;	2000US-0251989P.	PR			
XX	08-DEC-2000;	2000US-0251990P.	PR			
XX	11-DEC-2000;	2000US-0254097P.	PR			
XX	05-JAN-2001;	2001US-0259678P.	PR			
XX	(HUMA-) HUMAN GENOME SGT INC.					
PI	Rosen CA, Barash SC, Ruben SM;					
XX	WPI; 2001-476223/51.					
XX	Novel isolated prostate gland related polypeptide useful for diagnosis					
PT	and treatment of disorders of prostate such as prostatodystonia,					
PT	prostatosis, prostatitis, benign prostatic hypertrophy and malacoplaxia.					
PS	Claim 1; SEQ ID NO 338; 512pp; English.					
XX	The invention relates to novel isolated prostate gland related nucleic					
CC	acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,					
CC	prognosis, prevention, and/or treatment of diseases and/or disorders of					
CC	the prostate such as acute non-bacterial prostatitis, chronic non-					
CC	bacterial prostatitis, acute bacterial prostatitis, prostatodytonia,					
CC	prostatosis, granulomatous prostatitis, malacoplaxia, benign prostatic					
CC	hypertrophy or hyperplasia, and prostate neoplastic disorders, including					
CC	adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and					
CC	squamous cell carcinomas. (I), (II) and antibody to (II) are useful for					
CC	diagnosing and treating reproductive system disorders (Paget's disease),					
CC	autoimmune disorders (systemic lupus erythematosus, rheumatoid					
CC	arthritis), blood-related disorders (sickle cell anaemia),					
CC	hyperproliferative disorders, urinary system disorders					
CC	(glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory					
CC	disorders, musculoskeletal system disorders, neural activity and					
CC	neurological disorders (Alzheimer's disease and Parkinson's disease),					
CC	endocrine disorders (Addison's disease), gastrointestinal disorders					
CC	(inflammatory disorders), liver disorders (biliary liver cirrhosis),					
CC	pancreatic and gall bladder disorders, disorders of the large intestine,					
CC	developmental and inherited disorders, diseases at the cellular level,					
CC	and wound healing and epithelial cell proliferation. (I) or (II) is					
CC	useful to prevent skin aging, for preventing hair loss, to maintain					
CC	organs before transplantation, and as food additive or preservative.					
Query Match 83.3%; Score 15; DB 4; Length 3167;						
Best Local Similarity 100.0%; Pred. No. 99;						
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
OY	4 TGTTCGTGACTTT 18					
DB	1739 TGTTCGTGCTGACTTT 1753					

RESULT 12
AAL06260
ID AAL06260 standard; DNA; 3167 BP.
XX
AC AAL06260;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 8948.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN MO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.


```
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-46570/50.
DR
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 8948; 1297bp + Sequence Listing; English.
PS
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
SQ Sequence 3167 BP; 892 A; 608 C; 749 G; 918 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 4; Length 3167;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 TGTTCTGCTGACTTT 18
Db 1739 TGTTCTGCTGACTTT 1753
XX
RESULT 13
ABLO9088/c
ID ABL09088 standard; cDNA; 3190 BP.
XX
XX ABL09088;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 21746.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009221.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
PR
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX
XX P-PSDB; ABB64985.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
```

```
PT Interactions.
XX
PS Claim 1; SEQ ID NO 21746; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABJ30511), expressed DNA
CC sequences (ABL10840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3190 BP; 967 A; 700 C; 674 G; 849 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 4; Length 3190;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTGTTCTGCTGAC 15
Db 2024 TGCTGTTCTGCTGAC 2010
XX
RESULT 14
ADFS9752/c
ID ADF59752 standard; cDNA; 5919 BP.
XX
XX ADF59752;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Human contig polynucleotide sequence SEQ ID NO:2119.
DE
XX
XX biological activity; genetic engineering; hybridisation probe; oligomer;
KM primer; chromosome mapping; gene mapping; recombinant protein production;
KW human; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO2003080795-A2.
PN
XX
XX 02-OCT-2003.
PD
XX
XX 09-AUG-2002; 2002WO-US025485.
PF
XX
XX 09-AUG-2001; 2001US-0311261P.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
PI
XX
XX WPI; 2003-876918/81.
DR
XX
XX P-PSDB; ADF60204.
DR
XX
XX New polynucleotides, useful as hybridization probes, oligomers or
PT primers, for chromosome or gene mapping, for the recombinant production
PT of proteins, and for generating antisense DNA or RNA.
XX
XX Example 2; SEQ ID NO 2119; 571bp; English.
PS
XX
XX The present invention describes isolated polynucleotide sequences (I),
XX which encode polypeptides (II) with biological activity. Also described:
XX (1) a vector comprising (I); (2) an expression vector comprising (I); (3)
XX a host cell genetically engineered to comprise (I) which is operatively
XX associated with a regulatory sequence that modulates expression of (I) in
XX the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition
XX comprising the polypeptide of (4) and a carrier; (6) an antibody directed
XX against the polypeptide of (4); (7) detecting (I) or the polypeptide of
XX (4) in a sample; (8) identifying a compound that binds to the polypeptide
XX of (4); (9) producing the polypeptide of (4); and (10) a collection of
```

CC polynucleotides comprising at least one of the polynucleotide sequences
CC (1). The polynucleotides (1) can be used as hybridisation probes,
CC oligomers or primers, for chromosome or gene mapping, for the recombinant
CC production of proteins, and for generating antisense DNA or RNA. The
CC present sequence represents a human contig polynucleotide sequence, which
CC is used in an example from the present invention.

XX Sequence 5919 BP; 1846 A; 1292 C; 1202 G; 1579 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 10; Length 5919;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TGTTCGTGACTTT 18

Db 5871 TGTTCGTGACTTT 5857

RESULT 15

AAK6202
ID AAK6202 standard; DNA; 41335 BP.

XX AAK6202;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41014.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 14-AUG-2000; 2000US-0225759P.

XX PR 18-AUG-2000; 2000US-0226279P.

XX PR 22-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0226686P.

XX PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232357P.

PR 14-SEP-2000; 2000US-0232358P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236357P.

PR 29-SEP-2000; 2000US-0236358P.

PR 29-SEP-2000; 2000US-0236359P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239355P.

PR 13-OCT-2000; 2000US-0239357P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246529P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Barash SC, Ruben SM;
XX DR WPI; 2001-483426/52.
XX PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PS Disclosure; SEQ ID NO 41014; 3071bp + Sequence Listing; English.
XX CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytostatic
XX CC activity, and can be used in gene therapy and vaccine production. (I)
XX CC proteins and polynucleotides may be used in the prevention, diagnosis and
XX CC treatment of diseases associated with inappropriate (I) expression. For
XX CC example, they may be used to treat disorders associated with decreased
XX CC expression by rectifying mutations or deletions in a patient's genome
XX CC that affect the activity of (I) by expressing inactive proteins or to
XX CC supplement the patient's own production of (I). Additionally, (I)
XX CC polynucleotides may be used to produce the secreted (I), by inserting the
XX CC nucleic acids into a host cell and culturing the cell to express the
XX CC protein. (I) proteins and polynucleotides may be used to prevent,
XX CC diagnose and treat immune/haematopoietic-related diseases, especially
XX CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX CC to AAK87694 represent human immune/haematopoietic antigen genomic
XX CC sequences from the present invention. AAK54942 to AAK54950 and AAM62169
XX CC represent sequences used in the exemplification of the present invention
SQ Sequence 41335 BP; 9383 A; 10738 C; 10251 G; 10957 T; 0 U; 6 Other;

Query Match 83.3%; Score 15; DB 4; Length 41335;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCTGTTCTGCTGACT 16
DB 34054 GCTGTTCTGCTGACT 34068

This Page Blank (uspto)

121 GTTACTAGATCGGGATATCCCGGGAAATCGGTAACAAGCTTTTAAATAGTAAAA 180
181 GAGTAAATTTCACTTTGGGCACTTTTATACCGATTTTATACCACTTTTA 240
181 GAGTAAATTTCACTTTGGGCACTTTTATACCGATTTTATACCACTTTTA 240
241 ACTGATGTTTTCACCTTTGACCAAGTAACTTAACTTTGTTTATTTGACTATCCGA 300
241 ACTGATGTTTTCACCTTTGACCAAGTAACTTAACTTTGTTTATTTGACTATCCGA 300
301 CTCTCTTTCAGACATATGATATACCTTCAGTAAGCTTGTAAACGGGCGCCCTAGGA 360
301 CTCTCTTTCAGACATATGATATACCTTCAGTAAGCTTGTAAACGGGCGCCCTAGGA 360
361 TATCAAGCTTGTACACGACGACACTTTCACCTAGTGTGAGTGAATCCTGTTATC 420
361 TATCAAGCTTGTACACGACGACACTTTCACCTAGTGTGAGTGAATCCTGTTATC 420
421 TCTTCTGACACATACAGATATATTTATTTGATCATTTGAATCGTTTATTTCTTGA 480
421 TCTTCTGACACATACAGATATATTTATTTGATCATTTGAATCGTTTATTTCTTGA 480
481 AGCGGTTTCAATTTTTTTTACAGACGCTTTTTTATAGAGTGAACATCATATGCGGC 540
481 AGCGGTTTCAATTTTTTTTACAGACGCTTTTTTATAGAGTGAACATCATATGCGGC 540
541 ATAGGTGTATCATGCGGTATACCACTTAAACGATACCACTTTTATAGCAATGCGTGA 600
541 ATAGGTGTATCATGCGGTATACCACTTAAACGATACCACTTTTATAGCAATGCGTGA 600
601 GCGGCATCTCTTCCGCTACAGACACTTTTACATTAATTCTGCTGTTCAAAACCACT 660
601 GCGGCATCTCTTCCGCTACAGACACTTTTACATTAATTCTGCTGTTCAAAACCACT 660
661 GTAGGATAGCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
661 GTAGGATAGCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
721 ATAAATATGACAAACAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
721 ATAAATATGACAAACAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
781 TTCAACACGCTGCGACACGATGCAACGACCTTCCAAAGCAATATCTGAGGCGGACC 840
781 TTCAACACGCTGCGACACGATGCAACGACCTTCCAAAGCAATATCTGAGGCGGACC 840
841 GACGACGTAAGCAGGCGGTGCGATACGACGCTGCGGCAATCACTTCTTCCCTT 900
841 GACGACGTAAGCAGGCGGTGCGATACGACGCTGCGGCAATCACTTCTTCCCTT 900
901 TCTCTGCTTCAACTTGGGCGGCGAGTCTGCTAGACCCAGGAGATGCTGTGAGAGGA 960
901 TCTCTGCTTCAACTTGGGCGGCGAGTCTGCTAGACCCAGGAGATGCTGTGAGAGGA 960
961 GGTGCGGCGGCGCAATTTTATAGCTGAGCGAGACGAGCTTGGCCGAACTGACGA 1020
961 GGTGCGGCGGCGCAATTTTATAGCTGAGCGAGACGAGCTTGGCCGAACTGACGA 1020
1021 GCTCTGCGCAATATCAAGAAACCACTGAGGCGCTGCGGCTTACCCACCGCAGAGC 1080
1021 GCTCTGCGCAATATCAAGAAACCACTGAGGCGCTGCGGCTTACCCACCGCAGAGC 1080
1081 GGGGCTTGTGCGAGCGGTAGCTGCGGAGGAGGAGCGCGCTAGGGGGGCGCATGCTC 1140
1081 GGGGCTTGTGCGAGCGGTAGCTGCGGAGGAGGAGCGCGCTAGGGGGGCGCATGCTC 1140
1141 CAGCGCCCAAGAGAAAAAGAAAGAAAGGCGGAGATGATG 1183
1141 CAGCGCCCAAGAGAAAAAGAAAGAAAGGCGGAGATGATG 1183

RESULT 2
US-09-441-340-27/c

Sequence 27, Application US/09441340
Patent No. 6448476
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
TITLE OF INVENTION: Phosphonate Metabolizing Plants
FILE REFERENCE: 38-21(15303)
CURRENT FILING DATE: US/09/441,340
EARLIER APPLICATION NUMBER: 60/108,763
EARLIER FILING DATE: 1998-11-17
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 27
LENGTH: 2378
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: expression
OTHER INFORMATION: cassette comprising a plant promoter linked to an
OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
OTHER INFORMATION: transferase, and a termination sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (28) .. (965)
NAME/KEY: intron
LOCATION: (966) .. (1423)
NAME/KEY: transit peptide
LOCATION: (1440) .. (1667)
FEATURE:
NAME/KEY: CDS
LOCATION: (1668) .. (2099)
FEATURE:
NAME/KEY: terminator
LOCATION: (2114) .. (2369)
US-09-441-340-27

Query Match 17.1%; Score 202; DB 3; Length 2378;
Best Local Similarity 100.0%; Pred. No. 2.5e-89;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTAAATATGTAAGAAAGATTAATTTCACTTTGGGCGCACTTTTATACGATATTTT 222
DB 202 TTTTAAATATGTAAGAAAGATTAATTTCACTTTGGGCGCACTTTTATACGATATTTT 143
QY 223 ACTTATACCACTTTTAACTGATGTTTCACTTTGACGAGTAATCTTACCTTGT 282
DB 142 ACTTATACCACTTTTAACTGATGTTTCACTTTGACGAGTAATCTTACCTTGT 83
QY 283 TATTTGACTATCCGACTCTTCTTCAAGCATATGAATGACCTGAGTAAGCTTGT 342
DB 82 TATTTGACTATCCGACTCTTCTTCAAGCATATGAATGACCTGAGTAAGCTTGT 23
QY 343 ACGGCGCGCCCTAGGATATC 364
DB 22 ACGGCGCGCCCTAGGATATC 1

RESULT 3
US-09-068-101-5/c
Sequence 5, Application US/09068101
Patent No. 642960
GENERAL INFORMATION:
APPLICANT: PLANT GENETIC SYSTEMS N.V.
TITLE OF INVENTION: Improved Barstar Gene
FILE REFERENCE: 2121-1139P
CURRENT FILING DATE: US/09/068,101
EARLIER APPLICATION NUMBER: EP 96202446.9
EARLIER FILING DATE: 1996-09-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0

This Page Blank (uspto)

Qy 61 GATGAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120
Db 3022 GATGAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 3081
Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
Db 3082 GTTACTAGATCGGGGATATCCCGGGG 3108

RESULT 9
US-09-186-002-15
Sequence 15, Application US/09186002B
Patent No. 6489542
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
APPLICANT: Corbin, David R.
TITLE OF INVENTION: Improved Method for Transforming Plants to Express
FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002
CURRENT FILING DATE: 1998-11-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 15
LENGTH: 10252
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: unsure
LOCATION: (3660)..(3773)
OTHER INFORMATION: "n" = g, a, c, or t
FEATURE:
NAME/KEY: unsure
LOCATION: (4355)..(4407)
OTHER INFORMATION: "n" = g, a, c, or t
US-09-186-002-15

Query Match 12.4%; Score 147; DB 4; Length 10252;
Best Local Similarity 100.0%; Pred. No. 2.9e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
Db 2965 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 3024
Qy 61 GATGAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120
Db 3025 GATGAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 3084
Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
Db 3085 GTTACTAGATCGGGGATATCCCGGGG 3111

RESULT 10
US-09-186-002-13
Sequence 13, Application US/09186002B
Patent No. 6489542
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Method for Transforming Plants to Express
FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002
CURRENT FILING DATE: 1998-11-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13
LENGTH: 10339
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
NAME/KEY: unsure
LOCATION: (3687)..(3760)
OTHER INFORMATION: "n"=g, a, c, or t
FEATURE:
NAME/KEY: unsure
LOCATION: (4382)..(4434)
OTHER INFORMATION: "n" = g, a, c, or t
US-09-186-002-13

Query Match 12.4%; Score 147; DB 4; Length 10339;
Best Local Similarity 100.0%; Pred. No. 2.9e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
Db 3052 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 3111
Qy 61 GATGAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120
Db 3112 GATGAAACAAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 3171
Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
Db 3172 GTTACTAGATCGGGGATATCCCGGGG 3198

RESULT 11
US-09-441-340-27
Sequence 27, Application US/09441340
Patent No. 6489542
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
TITLE OF INVENTION: Phosphonate Metabolizing Plants
FILE REFERENCE: 38-21(15303)
CURRENT FILING DATE: 1998-11-16
EARLIER FILING DATE: 1998-11-17
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 27
LENGTH: 2378
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:expression
OTHER INFORMATION: cassette comprising a plant promoter linked to an
OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
OTHER INFORMATION: transferase, and a termination sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (28)..(965)
FEATURE:
NAME/KEY: intron
LOCATION: (966)..(1423)
FEATURE:
NAME/KEY: transic peptide
LOCATION: (1440)..(1667)
FEATURE:
NAME/KEY: CDS
LOCATION: (1668)..(2099)
FEATURE:
NAME/KEY: terminator
LOCATION: (2114)..(2369)
US-09-441-340-27

Query Match 11.8%; Score 140; DB 3; Length 2378;
Best Local Similarity 100.0%; Pred. No. 8.5e-59;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
|||||

This Page Blank (uspto)

Mon Feb 14 10:13:01 2005

us-10-790-430-8.01igo.rml

Page 6

Db 2239 GACGTTATTATGAGATGGGTTTATGATAGATCCGCAATTATATACATTAAATACGC 2298
QY 61 GATAGAAAACAAATATAGCGCCAACTAGATTAATTTTCGGCGGCGTGCATCTAT 120
Db 2299 GATAGAAAACAAATATAGCGCCAACTAGATTAATTTTCGGCGGCGTGCATCTAT 2358
QY 121 GTTACTAGATCGGGGATATC 140
Db 2359 GTTACTAGATCGGGGATATC 2378

RESULT 12
US-09-441-340-31
Sequence 31, Application US/09441340
Patent No. 6448476
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
TITLE OF INVENTION: Phosphonate Metabolizing Plants
FILE REFERENCE: 38-21(15303)
CURRENT APPLICATION NUMBER: US/09/441,340
EARLIER FILING DATE: 1999-11-16
EARLIER APPLICATION NUMBER: 60/108,763
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 31
LENGTH: 2436
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: monocot
OTHER INFORMATION: expression cassette comprising plant operable
OTHER INFORMATION: promoter linked to an intron, a sequence coding
OTHER INFORMATION: for an AMPA acetyltransferase, and a termination
OTHER INFORMATION: sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (26)..(640)
FEATURE:
NAME/KEY: intron
LOCATION: (670)..(1473)
FEATURE:
NAME/KEY: transit peptide
LOCATION: (1498)..(1725)
FEATURE:
NAME/KEY: CDS
LOCATION: (1726)..(2157)
FEATURE:
NAME/KEY: terminator
LOCATION: (2172)..(2427)
US-09-441-340-31

Query Match 11.8%; Score 140; DB 3; Length 2436;
Best Local Similarity 100.0%; Pred. No. 8.5e-59;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATAGATCCGCAATTATATACATTAAATACGC 60
Db 2297 GACGTTATTATGAGATGGGTTTATGATAGATCCGCAATTATATACATTAAATACGC 2356
QY 61 GATAGAAAACAAATATAGCGCCAACTAGATTAATTTTCGGCGGCGTGCATCTAT 120
Db 2357 GATAGAAAACAAATATAGCGCCAACTAGATTAATTTTCGGCGGCGTGCATCTAT 2416
QY 121 GTTACTAGATCGGGGATATC 140
Db 2417 GTTACTAGATCGGGGATATC 2436

RESULT 13
US-09-441-340-23
Sequence 23, Application US/09441340
Patent No. 6448476

GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
TITLE OF INVENTION: Phosphonate Metabolizing Plants
FILE REFERENCE: 38-21(15303)
CURRENT APPLICATION NUMBER: US/09/441,340
EARLIER FILING DATE: 1999-11-16
EARLIER APPLICATION NUMBER: 60/108,763
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 1630
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: expression
OTHER INFORMATION: cassette comprising plant operable promoter linked
OTHER INFORMATION: to a coding sequence encoding an AMPA
OTHER INFORMATION: acetyltransferase linked to a transcription
OTHER INFORMATION: termination sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (33)..(605)
FEATURE:
NAME/KEY: transit peptide
LOCATION: (627)..(892)
FEATURE:
NAME/KEY: CDS
LOCATION: (893)..(1324)
FEATURE:
NAME/KEY: terminator
LOCATION: (1350)..(1605)
US-09-441-340-23

Query Match 11.6%; Score 137; DB 3; Length 1630;
Best Local Similarity 100.0%; Pred. No. 2.6e-57;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATAGATCCGCAATTATATACATTAAATACGC 60
Db 1476 GACGTTATTATGAGATGGGTTTATGATAGATCCGCAATTATATACATTAAATACGC 1535
QY 61 GATAGAAAACAAATATAGCGCCAACTAGATTAATTTTCGGCGGCGTGCATCTAT 120
Db 1536 GATAGAAAACAAATATAGCGCCAACTAGATTAATTTTCGGCGGCGTGCATCTAT 1595
QY 121 GTTACTAGATCGGGGAT 137
Db 1596 GTTACTAGATCGGGGAT 1612

RESULT 14
US-09-098-219B-5
Sequence 5, Application US/09098219B
Patent No. 6441277

GENERAL INFORMATION:
APPLICANT: Barry, Gerard
APPLICANT: Cheikh, No. 6441277dine
TITLE OF INVENTION: Expression of Fructose 1,6 Biphosphosphate
TITLE OF INVENTION: Algalase in Transgenic Plants
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 121.68 Seconds
(Without alignments)
5630.828 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 18

Sequence: 1 taccacgcgacacacttc 18

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19023134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hlc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g981.*
9: gb_g982.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.9	1441	9	AG076194	AG076194 Pan trogl
2	83.3	247	9	CL445530	CL445530 ZMMBBD046
3	83.3	364	5	BU067975	BU067975 1632_E05
4	83.3	370	5	BU068295	BU068295 2556_F04
5	83.3	399	5	BU068642	BU068642 2562_H04
6	83.3	414	5	BU068903	BU068903 2569_E09
7	83.3	440	6	CB039236	CB039236 Ag_L3_31G
8	83.3	449	6	A2141658	A2141658 SP_0003_B
9	83.3	452	1	AL924785	AL924785
10	83.3	453	5	BU068182	BU068182_2555_B04
11	83.3	456	7	CK989655	CK989655 BgHC-2.38
12	83.3	459	6	CA954936	CA954936 k132e04.y
13	83.3	477	4	B1593866	B1593866 Ag_L3_06A
14	83.3	481	5	BU067587	BU067587 1626_F12
15	83.3	483	5	CA953980	CA953980 k137a10.y
16	83.3	498	4	BM515610	BM515610 k137e03.y
17	83.3	503	4	BM515613	BM515613 k137e03.y
18	83.3	508	4	BM515619	BM515619 k137e03.y
19	83.3	509	6	BU066631	BU066631 1609_B02
20	83.3	517	6	CB039563	CB039563 Ag_L3_37H
21	83.3	524	4	BM283870	BM283870 k128f01.y
22	83.3	524	4	BM568782	BM568782 k141h07.y
23	83.3	524	6	CA849289	CA849289 k110c10.y
24	83.3	524	6	CA953971	CA953971 k124h07.y

C 25	15	83.3	524	6	CA954253	CA954253 k140e12.y
C 26	15	83.3	524	6	CA954961	CA954961 k132h07.y
C 27	15	83.3	524	6	CA954983	CA954983 k134c02.y
C 28	15	83.3	527	2	BE604361	BE604361 WHE1405-1
C 29	15	83.3	537	6	CA954745	CA954745 k130a03.y
C 30	15	83.3	540	6	CA953797	CA953797 k122c05.y
C 31	15	83.3	546	6	BM283807	BM283807 k152g03.y
C 32	15	83.3	547	6	CD883084	CD883084 F1.112E01
C 33	15	83.3	549	5	BU067422	BU067422 1620_G02
C 34	15	83.3	549	6	CA954309	CA954309 k141c06.y
C 35	15	83.3	560	6	BM281999	BM281999 k124d01.y
C 36	15	83.3	560	6	CA849866	CA849866 k106d02.y
C 37	15	83.3	562	4	BM280788	BM280788 k105a07.y
C 38	15	83.3	562	6	CA954933	CA954933 k132d12.y
C 39	15	83.3	567	4	BM568585	BM568585 k139b05.y
C 40	15	83.3	569	6	CA953755	CA953755 k121e11.y
C 41	15	83.3	570	4	BM567304	BM567304 k101b03.y
C 42	15	83.3	571	6	CA850315	CA850315 k126g05.y
C 43	15	83.3	572	4	BM569276	BM569276 k146f07.y
C 44	15	83.3	572	6	CA850345	CA850345 k127b05.y
C 45	15	83.3	572	6	CA850345	CA850345 k127b05.y

ALIGNMENTS

RESULT 1	AG076194	1441 bp	DNA	linear	GSS 03-NOV-2001
LOCUS	Pan troglodytes DNA, clone: PTB-070F05.R, genomic survey sequence.				
DEFINITION	AG076194				
ACCESSION	AG076194.1	GI:16627996			
VERSION					
KEYWORDS	GSS.				
SOURCE	Pan troglodytes (chimpanzee)				
ORGANISM	Pan troglodytes				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.				
AUTHORS	1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	BAC end sequences of library PTB				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1441)				
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)				
REFERENCE	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.				
PRIMERS					
LIBRARY	Sequencing: M13Rev				
VECTOR	Vector : pKS145				
R.Site 1	: SacI				
R.Site 2	: SacI				
FEATURES	Location/Qualifiers				
source	1..1441				
	/organism="Pan troglodytes"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9598"				
	/clone="PTB-070F05.R"				
	/sex="male"				
	/cell_type="lymphoblast"				
	/clone_lib="PTB Chimpanzee Male BAC Library"				
ORIGIN					
Query Match	88.9%; Score 16; DB 9; Length 1441;				
Best local Similarity	100.0%; Pred. No. 14;				

Matches	16; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	TACGACGGACACT	16					
Db	1136	TACGACGGACACT	1151					

RESULT	2
CL445530/c	
LOCUS	247 bp DNA linear GSS 24-MAR-2006
DEFINITION	ZMMBBBD0463D18f ZMMBBB (HindIII) zea mays genomic clone
	ZMMBBBD0463D18 5', genomic survey sequence.

REFERENCE	1 (bases 1 to 247)
AUTHORS	Bharti,A.K., Young,S., Kavchok,S., Kelzer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE	Sequencing of the maize genome at PGR (2003c)
JOURNAL	Unpublished (2003)
COMMENT	Contact: Bharti,A.K.

FEATURES	Location/Qualifiers
source	1. .247

ORIGIN

Query Match	83.3%	Score 15;	DB 9;	Length 247;
Best Local Similarity	100.0%;	Pred. No. 68;		
Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

RESULT 3		
BU067975	364 bp	mRNA linear EST 27-AUG-2002
LOCUS	BU067975	
DEFINITION	1632_B05_J10ZT5 Mature perithecia <i>Gibberella zeae</i> cDNA, mRNA	

RESULT 5
 BU068642 399 bp mRNA linear EST 27-AUG-2002
 LOCUS BU068642
 DEFINITION 2562_H04_P08275 Mature perithecia *Gibberella zeae* cDNA, mRNA
 sequence.
 ACCESSION BU068642
 VERSION BU068642.1 GI:22509831
 KEYWORDS EST.
 SOURCE *Gibberella zeae*
 ORGANISM *Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocerales; Nectriaceae; Gibberella.*
 REFERENCE 1 (bases 1 to 399)
 AUTHORS Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
 TITLE Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*)
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)
 MEDLINE 22508120
 PUBMED 12620255
 COMMENT Contact: Frances Trail
 Department of Plant Biology
 Michigan State University
 East Lansing, MI 48824, USA
 Tel: 517 432 2939
 Fax: 517 353 1926
 Email: trail@msu.edu.
 Location/Qualifiers
 1..399
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="taxon:5518"
 /clone_id="Mature perithecia"
 /note="Vector: Ziplox; Site_1: NotI; Site_2: SalI"

ORIGIN
 Query Match 83.3%; Score 15; DB 5; Length 399;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18
 |||||
 Db 300 CACGCGACACACTTC 314

RESULT 6
 BU068903 414 bp mRNA linear EST 27-AUG-2002
 LOCUS BU068903
 DEFINITION 2569_E09_J17275 Mature perithecia *Gibberella zeae* cDNA, mRNA
 sequence.
 ACCESSION BU068903
 VERSION BU068903.1 GI:22510092
 KEYWORDS EST.
 SOURCE *Gibberella zeae*
 ORGANISM *Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocerales; Nectriaceae; Gibberella.*
 REFERENCE 1 (bases 1 to 414)
 AUTHORS Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
 TITLE Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*)
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)
 MEDLINE 22508120
 PUBMED 12620255
 COMMENT Contact: Frances Trail
 Department of Plant Biology
 Michigan State University
 East Lansing, MI 48824, USA
 Tel: 517 432 2939
 Fax: 517 353 1926
 Email: trail@msu.edu.
 Location/Qualifiers

FEATURES
 source
 1..414
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="taxon:5518"
 /clone_id="Mature perithecia"
 /note="Vector: Ziplox; Site_1: NotI; Site_2: SalI"

ORIGIN
 Query Match 83.3%; Score 15; DB 6; Length 440;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

source
 1..414
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="taxon:5518"
 /clone_id="Mature perithecia"
 /note="Vector: Ziplox; Site_1: NotI; Site_2: SalI"

ORIGIN
 Query Match 83.3%; Score 15; DB 5; Length 414;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18
 |||||
 Db 299 CACGCGACACACTTC 313

RESULT 7
 CB039236/c 440 bp mRNA linear EST 15-JAN-2003
 LOCUS CB039236
 DEFINITION As L3_31G08 SKPL *Ascaris suum* (parasitic nematode) L3 stage *Ascaris suum* cDNA clone As L3_31G08 5' similar to p41105 60S ribosomal protein L28. Mus musculus (Mouse), mRNA sequence.
 ACCESSION CB039236
 VERSION CB039236.1 GI:27758480
 KEYWORDS EST.
 SOURCE *Ascaris suum* (pig roundworm)
 ORGANISM *Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascarididae; Ascaris.*
 REFERENCE 1 (bases 1 to 440)
 AUTHORS Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guiliano, D., Hall, N., Quayle, M. and Bartrell, B.
 TITLE Edinburgh University/Sanger Centre Nematode EST Project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Blaxter ML
 Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared for Malcolm Kennedy by Joyce Moore,
 Glasgow University. Sequencing was performed by the Pathogen
 Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike
 Quail & Bart Bartrell).
 PCR Primers
 FORWARD: T3
 BACKWARD: T7PL
 Plate: 31 row: G column: 08
 Seq primer: SKPL
 High quality sequence stop: 296.
 Location/Qualifiers
 1..440
 /organism="Ascaris suum"
 /mol_type="mRNA"
 /db_xref="taxon:6253"
 /clone="As L3_31G08"
 /sex="mixed"
 /dev_stage="L3"
 /clone_id="Ascaris suum (parasitic nematode) L3 stage"
 /note="Vector: Lambda ZAP II; Site_1: EcoRI (5'end); The
 Site_2: (3'end); *Ascaris suum* is a parasite of pigs. The
 library was constructed from mRNA from *Ascaris suum* L3
 stage."

ORIGIN
 Query Match 83.3%; Score 15; DB 6; Length 440;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACCGGACACACTTC 18
|||||
405 CACCGGACACACTTC 391

Db 405 CACCGGACACACTTC 391

RESULT 8
LOCUS A2141658 449 bp DNA linear GSS 28-AUG-2000
DEFINITION SP 0003 B1 A08 SP66 Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=3 Col=15 Row=B, genomic survey sequence.

ACCESSION A2141658
VERSION A2141658.1 GI:8293561
KEYWORDS GSS.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryote; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidae; Euechinoidae; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 449)
Cameron, R.A., Mahitras, G., Raat, J.P., Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray, G.A., Bittensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H. and Hood, L.
A sea urchin genome project: Sequence scan, virtual map, and additional resources
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
20402566
10920195
COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 3 row: B column: 15
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 449.

FEATURES
source location/Qualifiers
1..449
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone="Plate=3 Col=15 Row=B"
/clone_lib="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACs.6; BAC Clones in E-Coli DH10B"

ORIGIN
Query Match 83.3%; Score 15; DB 8; Length 449;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACCGGACACACTTC 18
|||||
23 CACCGGACACACTTC 37

Db 23 CACCGGACACACTTC 37

RESULT 9
LOCUS AL924785 452 bp mRNA linear EST 06-JUL-2004
DEFINITION AL924785 PUR-Z1+Z2 Danio rerio cDNA clone 145-H10-2, mRNA sequence.
ACCESSION AL924785
VERSION AL924785.1 GI:23191365
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 452)
AUTHORS Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Sun, A., He, Y., Ma, W., Wang, W., Wen, Z. and Peng, J.
15000 unique zebrafish EST clusters and their future use in microarray for profiling gene expression patterns during embryogenesis
Genome Res. 13 (3), 455-466 (2003)

JOURNAL MEDLINE 12618376
PUBMED 22505427
COMMENT Contact: Peng J
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengj@imcb.a-star.edu.sg
Clone requests: info@openbioystems.com
Open Bioystems,
6705 Odyssey Drive, Huntsville, AL 35806.

FEATURES
source location/Qualifiers
1..452
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="145-H10-2"
/issue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_lib="PJR-Z1+Z2"

ORIGIN
Query Match 83.3%; Score 15; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACCGGACACACTTC 18
|||||
26 CACCGGACACACTTC 12

Db 26 CACCGGACACACTTC 12

RESULT 10
LOCUS BU068182 453 bp mRNA linear EST 27-AUG-2002
DEFINITION 2555_B04_C07ZT5 Mature perithecia Gibberella zeae cDNA, mRNA sequence.

ACCESSION BU068182
VERSION BU068182.1 GI:22509371
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

REFERENCE 1 (bases 1 to 453)
AUTHORS Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
Analysis of expressed sequence tags from Gibberella zeae (anamorph Fusarium graminearum)
Fungal Genet. Biol. 38 (2), 187-197 (2003)

JOURNAL MEDLINE 22508120
PUBMED 12620255
COMMENT Contact: Frances Trail
Department of Plant Biology
Michigan State University
East Lansing, MI 48824, USA
Tel: 517 432 2939
Fax: 517 353 1926
Email: trail@msu.edu.

FEATURES
source location/Qualifiers
1..453
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="NRRL 31084"
/db_xref="taxon:5518"
/clone_lib="Mature perithecia"
/note="Vector: Ziplox; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 83.3%; Score 15; DB 5; Length 453;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACTTC 18
 |||
 DB 299 CACGCGACACTTC 313

RESULT 11
 LOCUS CK989655 456 bp mRNA linear EST 19-MAR-2004
 DEFINITION BgHC-2.38 BgHC *Biomphalaria glabrata* cDNA 5', mRNA sequence.
 ACCESSION CK989655
 VERSION CK989655.1 GI:45596692
 KEYWORDS EST.
 SOURCE *Biomphalaria glabrata* (bloodfluke planorb)
 ORGANISM *Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora; Lymnaeidae; Planorbidae; Biomphalaria.*
 1 (bases 1 to 456)
 Mitter, G., Gallinder, R., Tisseyre, P., Allienne, J.-F., Carmignani, C. and Coureau, C.
 ESTs from *Biomphalaria glabrata* hemocytes
 Unpublished (2004)
 Contact: Mita Guillaume
 Parasitologie fonctionnelle et evolutive.
 CNRS / Univ. Perpignan
 52 Ave Paul Alduy, Universite, 66 860 Perpignan cedex, France
 Tel: 33 3 68 66 21 88
 Fax: 33 3 68 66 22 81
 Email: mita@univ-perp.fr
 Seg primer: CGAAGTATCTCGACG.
 Location/Qualifiers
 1. 456
 /organism="Biomphalaria glabrata"
 /mol_type="mRNA"
 /isolate="Bg. Bra strain"
 /db_xref="taxon:6526"
 /sex="hermaphrodite"
 /cell_type="circulating hemocytes"
 /clone_id="BgHC"
 /note="Vector: pDNR-LIB; The library was designed with 500 ng of total RNA using the creator SMART cDNA library construction kit (Clontech)."

ORIGIN
 Query Match 83.3%; Score 15; DB 7; Length 456;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCAGCGACACTT 17
 |||
 DB 184 CCAGCGACACTT 198

RESULT 12
 LOCUS CA954936 459 bp mRNA linear EST 02-JAN-2003
 DEFINITION K132e04.y1 *Ascaris suum* embryo SL1 TOPO v1 *Ascaris suum* cDNA 5' similar to SW:RL28_MOUSE P41105 60S RIBOSOMAL PROTEIN L28. ;, mRNA sequence.
 ACCESSION CA954936
 VERSION CA954936.1 GI:27471296
 KEYWORDS EST.
 SOURCE *Ascaris suum* (pig roundworm)
 ORGANISM *Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea; Ascarididae; Ascaris.*
 1 (bases 1 to 459)
 McCarter, J., Clifton, S., Chiappelli, B., Pape, D., Martin, J., Wylie, T., Dance, M., Mitter, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, B., Bennett, J., Franklin, C., Tsagaris, H., R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Pearson, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Embryo cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into PCR11-TOPO(Invitrogen) following the TOPO TA cloning protocol. 30-60 cell embryo material was provided by Dr. Richard Davis of City University of New York Graduate Center, College of Staten Island, Staten Island, NY (redavis@postbox.csi.cuny.edu).
 Seg primer: SL1 primer
 High quality sequence stop: 437.
 Location/Qualifiers
 1. 459
 /organism="Ascaris suum"
 /mol_type="mRNA"
 /db_xref="taxon:6253"
 /dev_stage="30-60 cell embryo"
 /lab_host="DH10B"
 /clone_id="Ascaris suum embryo SL1 TOPO v1"
 /note="Vector: PCR11-TOPO (Invitrogen); Site_1: EcoRI; Site_2: EcoRI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Embryo cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into PCR11-TOPO(Invitrogen) following the TOPO TA cloning protocol. 30-60 cell embryo material was provided by Dr. Richard Davis of City University of New York Graduate Center, College of Staten Island, Staten Island, NY (redavis@postbox.csi.cuny.edu)."

ORIGIN
 Query Match 83.3%; Score 15; DB 6; Length 459;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACTTC 18
 |||
 DB 425 CACGCGACACTTC 411

RESULT 13
 LOCUS B1593866 477 bp mRNA linear EST 07-SEP-2001
 DEFINITION Ae L3 06A09 SKPL *Ascaris suum* (parasitic nematode) L3 stage *Ascaris suum* cDNA clone Ae L3 06A09 5' similar to ref|NP_033107.1| ribosomal protein L28 - Mus musculus, mRNA sequence.
 ACCESSION B1593866
 VERSION B1593866.1 GI:15497353
 KEYWORDS EST.
 SOURCE *Ascaris suum* (pig roundworm)
 ORGANISM *Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea; Ascarididae; Ascaris.*
 1 (bases 1 to 477)
 Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guilianno, D., Hall, N., Quayle, M. and Barrett, B.
 Edinburgh University/Sanger Centre Nematode EST Project
 Unpublished (2000)

COMMENT

Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared for Malcolm Kennedy by Joyce Moore,
Glasgow University. Sequencing was performed by the Pathogen
Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike
Quail & Bart Barrell).

FEATURES

source
1. .477
/organism="Ascaris suum"
/mol_type="mRNA"
/db_xref="taxon:6253"
/clone="As_L3_06A09"
/sex="mixed"
/dev_stage="L3"
/clone_lib="Ascaris suum (parasitic nematode) L3 stage"
/note="Vector: Lambda ZAP II; Site 1: EcoRI (5' end); The
Site 2: (3' end); Ascaris suum is a parasite of pigs. The
library was constructed from mRNA from Ascaris suum L3
stage."

ORIGIN

Query Match 83.3%; Score 15; DB 4; Length 477;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4 CACGCGACACACTTC 18
406 CACGCGACACACTTC 392

RESULT 14

BU067587 481 bp mRNA linear EST 27-AUG-2002
LOCUS 1626_F12_K24ZT5 Mature perithecium Gibberella zeae cDNA, mRNA
DEFINITION
sequence.
ACCESSION
BU067587
VERSION
BU067587.1 GI:22508776
KEYWORDS
EST.
SOURCE
Gibberella zeae
ORGANISM
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

REFERENCE

1 (bases 1 to 481)
Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
Analysis of expressed sequence tags from Gibberella zeae (anamorph
Fusarium graminearum)
Fungal Genet. Biol. 38 (2), 187-197 (2003)

AUTHORS

Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.

TITLE

Analysis of expressed sequence tags from Gibberella zeae (anamorph
Fusarium graminearum)

JOURNAL

Fungal Genet. Biol. 38 (2), 187-197 (2003)

MEDLINE

22508120
Fungal Genet. Biol. 38 (2), 187-197 (2003)

PUBMED

12620255
Contact: Frances Trail
Department of Plant Biology
Michigan State University
East Lansing, MI 48824, USA
Tel: 517 432 2939
Fax: 517 353 1926
Email: trail@msu.edu.

FEATURES

source
1. .481
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="NRRL 31084"

ORIGIN

/db_xref="taxon:5518"
/clone_lib="Mature perithecium"
/note="Vector: Zipox; Site 1: NotI; Site 2: SalI"
Query Match 83.3%; Score 15; DB 5; Length 481;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4 CACGCGACACACTTC 18
301 CACGCGACACACTTC 315

RESULT 15

CA953980/c 483 bp mRNA linear EST 02-JAN-2003
LOCUS K137a10.y1 Ascaris suum embryo SL1 TOPO v1 Ascaris suum cDNA 5'
DEFINITION similar to SW:RL28_MOUSE P41105 60S RIBOSOMAL PROTEIN L28.1, mRNA
sequence.
ACCESSION
CA953980
VERSION
CA953980.1 GI:27470340
KEYWORDS
EST.
SOURCE
Ascaris suum (pig roundworm)
ORGANISM
Ascaris suum
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae;
Ascaridoidea; Ascarididae; Ascaris.

REFERENCE

AUTHORS

McCarter, J., Clifton, S., Chapelli, B., Page, D., Martin, J.,
Wyllie, T., Dante, M., Maria, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Franklin, C.,
Tsagarelashvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Pearson, B., Swaller, T.,
Harvey, N., Schurk, R., Kohn, S., Shih, T., Jackson, Y., Cardenas, M.,
McCam, R., Waterston, R. and Wilson, R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

JOURNAL

The library was constructed by Claire Murphy and Dr. James McCarter
at Washington University, St. Louis. Oligo (dT)-SL1 PCR based
library. Embryo cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo (dT) on the 3' end were
non-directionally cloned into pCRIT-TOPO (Invitrogen) following the
TOPO TA cloning protocol. 30-60 cell embryo material was provided
by Dr. Richard Davis of City University of New York Graduate
Center, College of Staten Island, Staten Island, NY
(redavis@postbox.csi.cuny.edu).
Putative full length read
The vector to vector length is 588
Seq primer: SL1 primer.

FEATURES

source

1. .483
/organism="Ascaris suum"
/mol_type="mRNA"
/db_xref="taxon:6253"
/dev_stage="30-60 cell embryo"
/lab_host="DH10B"
/clone_lib="Ascaris suum embryo SL1 TOPO v1"
/note="Vector: pCRIT-TOPO (Invitrogen); Site 1: EcoRI;
Site 2: EcoRI; The library was constructed by Claire
Murphy and Dr. James McCarter at Washington University,
St. Louis. Oligo (dT)-SL1 PCR based library. Embryo cDNA
PCR products of size >400 nucleotides containing SL1 on
the 5' end and oligo (dT) on the 3' end were
non-directionally cloned into pCRIT-TOPO (Invitrogen)
following the TOPO TA cloning protocol. 30-60 cell embryo
material was provided by Dr. Richard Davis of City

University of New York Graduate Center, College of Staten
Island, Staten Island, NY (redavis@postbox.csi.cuny.edu). "

ORIGIN

Query Match 83.3%; Score 15; DB 6; Length 483;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACTTC 18
|||||
Db 421 CACGCGACACTTC 407

Search completed: February 10, 2005, 17:01:56
Job time : 124.68 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 29.935 Seconds
(without alignments)
3460.943 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 1 taccacgcacacacttc 18

Scoring table: OLIGO NTC
Gapop 60.0 , Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	9	US-09-872-051-10
2	18	100.0	18	18	US-10-790-430-10
3	18	100.0	1183	9	US-09-872-051-8
4	18	100.0	1183	18	US-10-790-430-8
5	17	94.4	466	18	US-10-425-115-114310
6	16	88.9	1338	17	US-10-369-493-33178
7	15	83.3	687	18	US-10-653-047-1012
8	15	83.3	721	18	US-10-425-115-14852
9	15	83.3	747	18	US-10-781-014-753
10	15	83.3	3309400	9	US-09-738-626-1
11	14	77.8	435	11	US-09-864-408A-1183

C 12	147	77.8	435	18	US-10-357-930-31761	Sequence 31761, A
C 13	14	77.8	437	18	US-10-357-930-40730	Sequence 40730, A
C 14	14	77.8	435	18	US-10-357-930-1395	Sequence 1395, Ap
C 15	14	77.8	484	18	US-10-357-930-10564	Sequence 10564, A
C 16	14	77.8	732	18	US-10-363-345A-1539	Sequence 1539, Ap
C 17	14	77.8	732	18	US-10-363-345A-1540	Sequence 1540, Ap
C 18	13	72.2	314	9	US-09-867-550-279	Sequence 279, App
C 19	13	72.2	419	17	US-10-369-493-30016	Sequence 30016, A
C 20	13	72.2	505	18	US-10-425-115-114679	Sequence 114679, A
C 21	13	72.2	564	18	US-10-425-115-150679	Sequence 150679, A
C 22	13	72.2	583	18	US-10-425-115-75209	Sequence 75209, A
C 23	13	72.2	729	18	US-10-363-345A-31705	Sequence 31705, A
C 24	13	72.2	729	18	US-10-363-345A-31706	Sequence 31706, A
C 25	13	72.2	789	17	US-10-424-599-79579	Sequence 79579, A
C 26	13	72.2	830	17	US-10-425-114-33981	Sequence 33981, A
C 27	13	72.2	840	18	US-10-363-345A-22085	Sequence 22085, A
C 28	13	72.2	840	18	US-10-363-345A-22086	Sequence 22086, A
C 29	13	72.2	863	18	US-10-425-115-6466	Sequence 6466, Ap
C 30	13	72.2	904	18	US-10-363-345A-22235	Sequence 22235, A
C 31	13	72.2	904	18	US-10-363-345A-22236	Sequence 22236, A
C 32	13	72.2	1023	18	US-10-425-115-145387	Sequence 145387, A
C 33	13	72.2	1125	17	US-10-424-599-8880	Sequence 8880, Ap
C 34	13	72.2	1189	18	US-10-767-701-12212	Sequence 12212, A
C 35	13	72.2	1251	18	US-10-437-963-43431	Sequence 43431, A
C 36	13	72.2	1344	17	US-10-365-742-35	Sequence 35, Appl
C 37	13	72.2	1396	18	US-10-363-345A-40673	Sequence 40673, A
C 38	13	72.2	1396	18	US-10-363-345A-40674	Sequence 40674, A
C 39	13	72.2	1620	17	US-10-282-114-14808	Sequence 14408, A
C 40	13	72.2	1687	17	US-10-425-114-15589	Sequence 15589, A
C 41	13	72.2	1743	18	US-10-739-930-1844	Sequence 1844, Ap
C 42	13	72.2	1832	17	US-10-425-114-25898	Sequence 25898, A
C 43	13	72.2	2000	9	US-09-938-842A-4558	Sequence 4558, Ap
C 44	13	72.2	2000	11	US-09-938-842A-4558	Sequence 4558, Ap
C 45	13	72.2	2587	17	US-10-424-599-57089	Sequence 57089, A

ALIGNMENTS

RESULT 1
US-09-872-051-10
Sequence 10, Application US/09872051
Patent No. US20020033960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
TITLE OF INVENTION: Corn Event PV-ZMGT32(mk603) and Composition and Methods for Det.
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
NAME/KEY: source
LOCATION: (1)..(18)
OTHER INFORMATION: zea maize plastid DNA and vector DNA
US-09-872-051-10
Query Match 100.0%; Score 18; DB 9; Length 18;

Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18
DB 1 TACCACGCGACACTTC 18

RESULT 2

US-10-790-430-10
Sequence 10, Application US/10790430
Publication No. US20040139493A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/10/790,430
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(18)
OTHER INFORMATION: zea maize plastid DNA and vector DNA
US-10-790-430-10

Query Match 100.0%; Score 18; DB 18; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18
DB 1 TACCACGCGACACTTC 18

RESULT 3

US-09-872-051-8
Sequence 8, Application US/09872051
Patent No. US20020013960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rps4
OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

Query Match 100.0%; Score 18; DB 9; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18
DB 373 TACCACGCGACACTTC 390

RESULT 4

US-10-790-430-8
Sequence 8, Application US/10790430
Publication No. US20040139493A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/10/790,430
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rps4
OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-10-790-430-8

Query Match 100.0%; Score 18; DB 18; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18
DB 373 TACCACGCGACACTTC 390

RESULT 5

US-10-425-115-114310

```
/ Sequence 114310, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 114310
/ LENGTH: 466
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MFT4577_35734C.1
US-10-425-115-114310

Query Match          94.4%; Score 17; DB 18; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACCACGCGACACACTT 17
Db 267 TACCACGCGACACACTT 283

RESULT 6
US-10-369-493-33178/c
/ Sequence 33178, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 33178
/ LENGTH: 1338
/ TYPE: DNA
/ ORGANISM: Xylella fastidiosa
US-10-369-493-33178

Query Match          88.9%; Score 16; DB 17; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACCACGCGACACACTT 17
Db 513 ACCACGCGACACACTT 498

RESULT 7
US-10-653-047-1012
/ Sequence 1012, Application US/10653047
/ Publication No. US20040229367A1
/ GENERAL INFORMATION:
/ APPLICANT: Randy M. Berka
/ APPLICANT: Michael W. Rey
/ APPLICANT: Jeffrey R. Shuster
/ APPLICANT: Sakari Kauppinen
/ APPLICANT: Ib Groch Clausen
```

```
/ APPLICANT: Peter Bjarke Olsen
/ TITLE OF INVENTION: Methods For Monitoring Multiple Gene
/ FILE REFERENCE: 5849,200-US
/ CURRENT APPLICATION NUMBER: US/10/653,047
/ CURRENT FILING DATE: 2003-08-29
/ PRIOR APPLICATION NUMBER: US/09/533,559
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 09/273,623
/ PRIOR FILING DATE: 1999-03-22
/ NUMBER OF SEQ ID NOS: 7860
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1012
/ LENGTH: 687
/ TYPE: DNA
/ ORGANISM: Fusarium venenatum
US-10-653-047-1012

Query Match          83.3%; Score 15; DB 18; Length 687;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTTC 18
Db 531 CACGCGACACACTTC 545

RESULT 8
US-10-425-115-14852
/ Sequence 14852, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 14852
/ LENGTH: 721
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)-(721)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MFT4577_113531C.1
US-10-425-115-14852

Query Match          83.3%; Score 15; DB 18; Length 721;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTTC 18
Db 476 CACGCGACACACTTC 490

RESULT 9
US-10-781-014-753/c
/ Sequence 753, Application US/10781014
/ Publication No. US20040180408A1
/ GENERAL INFORMATION:
/ APPLICANT: Pompejus, Markus
/ APPLICANT: Krieger, Burkhard
/ APPLICANT: Schröder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Haberer, Gregor
```

```
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
/ TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
/ TITLE OF INVENTION: PRODUCTION
/ FILE REFERENCE: BGI-126CCN
/ CURRENT APPLICATION NUMBER: US/10/781,014
/ CURRENT FILING DATE: 2004-02-17
/ PRIOR APPLICATION NUMBER: US 09/602,740
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141,031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/143,208
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 60/151,572
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19931412.8
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931413.6
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931419.5
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931420.9
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931424.1
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931428.4
/ PRIOR FILING DATE: 1999-07-08
/ REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 784
/ SEQ ID NO 753
/ LENGTH: 747
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (101)..(724)
/ OTHER INFORMATION: RXN02554
/ US-10-781-014-753

Query Match      83.3%; Score 15; DB 18; Length 747;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 CACGCGACACTTC 18
Db      21 CACGCGACACTTC 7
```

RESULT 10

```
US-09-738-626-1/c
/ Sequence 1, Application US/09738626
/ Publication No. US20020197605A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIKO
/ APPLICANT: HAYASHI, MIKIRO
/ APPLICANT: OCHIAI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAOKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
```

```
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO 1
/ LENGTH: 3309400
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ US-09-738-626-1
```

```
Query Match      83.3%; Score 15; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 CACGCGACACTTC 18
Db      3028188 CACGCGACACTTC 3028174
```

RESULT 11

```
US-09-864-408A-1183
/ Sequence 1183, Application US/09864408A
/ Publication No. US20040009474A1
/ GENERAL INFORMATION:
/ APPLICANT: Leach, Martin D.
/ TITLE OF INVENTION: NO. US20040009474A1 Human Polynucleotides and Polypeptides Enc
/ FILE REFERENCE: 21402-012
/ CURRENT APPLICATION NUMBER: US/09/864,408A
/ CURRENT FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 60/206,690
/ PRIOR FILING DATE: 2000-05-24
/ NUMBER OF SEQ ID NOS: 9068
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1183
/ LENGTH: 435
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-864-408A-1183
```

```
Query Match      77.8%; Score 14; DB 11; Length 435;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 ACGCGACACTTC 18
Db      261 ACGCGACACTTC 274
```

RESULT 12

```
US-10-357-930-31761/c
/ Sequence 31761, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
```

;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 31761
;; LENGTH: 435
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-357-930-31761

Query Match 77.8%; Score 14; DB 18; Length 435;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTT 17
DB 286 CACGCGACACACTT 273

RESULT 13
US-10-357-930-40730/c
; Sequence 40730, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40730
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-40730

Query Match 77.8%; Score 14; DB 18; Length 435;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTT 17
DB 286 CACGCGACACACTT 273

RESULT 14
US-10-357-930-1395/c
; Sequence 1395, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF

;; TITLE OF INVENTION: HUMAN PROSTATE CANCER
;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; PRIOR FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: 09/785,276
;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1395
;; LENGTH: 437
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: 36, 86, 87
;; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-1395

Query Match 77.8%; Score 14; DB 18; Length 437;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTT 17
DB 288 CACGCGACACACTT 275

RESULT 15
US-10-357-930-10564/c
; Sequence 10564, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10564
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-357-930-10564

Query Match 77.8%; Score 14; DB 18; Length 484;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTT 17
 |||||
 Db 335 CACGCGACACACTT 322

Search completed: February 9, 2005, 11:12:45
 Job time : 34.935 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 31.5981 Seconds
(without alignments)
3460.943 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19
Sequence: 1 tctagcgccgacgcgtcg 19

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	9	US-09-872-051-9
2	19	100.0	19	18	US-10-790-430-9
3	19	100.0	498	9	US-09-872-051-7
4	19	100.0	498	18	US-10-790-430-7
5	16	84.2	1044	15	US-10-156-761-1684
6	16	84.2	35167	18	US-10-645-794-5
7	16	84.2	35167	18	US-10-645-883A-1
8	16	84.2	9025608	15	US-10-156-761-1
9	15	78.9	521	17	US-10-424-599-43339
10	15	78.9	797	18	US-10-767-701-9887
11	15	78.9	1339	18	US-10-437-963-26586

12	15	78.9	1398	18	US-10-437-963-33098	Sequence 33098, A
13	15	78.9	1780	18	US-10-473-390-1	Sequence 1, Appl1
14	15	78.9	5013	18	US-10-437-963-60106	Sequence 60106, A
15	15	78.9	14800	18	US-10-473-390-3	Sequence 3, Appl1
16	14	73.7	218	18	US-10-425-115-172757	Sequence 172757, A
17	14	73.7	226	18	US-10-425-115-47913	Sequence 47913, A
18	14	73.7	228	9	US-09-983-965-4908	Sequence 4908, Ap
19	14	73.7	228	17	US-10-369-493-32240	Sequence 32240, A
20	14	73.7	292	9	US-09-960-352-7737	Sequence 7737, Ap
21	14	73.7	328	17	US-10-424-599-79875	Sequence 79875, A
22	14	73.7	410	18	US-10-425-115-174641	Sequence 174641, A
23	14	73.7	414	10	US-09-918-995-17011	Sequence 17011, A
24	14	73.7	436	9	US-09-983-965-2307	Sequence 2307, Ap
25	14	73.7	465	16	US-10-029-386-25290	Sequence 25290, A
26	14	73.7	473	18	US-10-425-115-29397	Sequence 29397, A
27	14	73.7	500	9	US-09-754-997A-43171	Sequence 43, Appl
28	14	73.7	512	17	US-10-424-599-177151	Sequence 127151, A
29	14	73.7	549	18	US-10-437-963-25717	Sequence 25717, A
30	14	73.7	610	18	US-10-767-701-231177	Sequence 23117, A
31	14	73.7	619	18	US-10-767-795-5965	Sequence 5965, Ap
32	14	73.7	619	18	US-10-437-963-61015	Sequence 61015, A
33	14	73.7	629	18	US-10-425-115-164400	Sequence 164400, A
34	14	73.7	635	18	US-10-437-963-9105	Sequence 9105, Ap
35	14	73.7	642	18	US-10-437-963-90735	Sequence 90735, A
36	14	73.7	676	18	US-10-437-963-85113	Sequence 85113, A
37	14	73.7	754	18	US-10-437-963-85113	Sequence 824, App
38	14	73.7	834	17	US-10-260-238-824	Sequence 176769, A
39	14	73.7	848	18	US-10-425-115-176769	Sequence 4711, Ap
40	14	73.7	870	17	US-10-260-238-4711	Sequence 127149, A
41	14	73.7	960	17	US-10-424-599-127149	Sequence 14674, A
42	14	73.7	1146	17	US-10-282-122A-14674	Sequence 556, App
43	14	73.7	1209	10	US-09-826-509-556	Sequence 2, Appl1
44	14	73.7	1209	18	US-10-472-365-2	Sequence 556, App
45	14	73.7	1209	19	US-10-925-095-556	Sequence 556, App

ALIGNMENTS

RESULT 1
US-09-872-051-9
Sequence 9, Application US/09872051
Patent No. US20020013960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OR INVENTION: Corn Event PV-ZMGN32 (nke603) and Composition and Methods for Det
FILE REFERENCE: 38-21 (52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence
NAME/KEY: source
LOCATION: (1)..(19)
OTHER INFORMATION: Zea maize genomic and vector DNA
US-09-872-051-9
Query Match 100.0%; Score 19; DB 9; Length 19;

Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19
Db 1 TGTAGCGGCCACGCGTGG 19

RESULT 2

US-10-790-430-9
Sequence 9, Application US/10790430
Publication No. US20040139493A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/10/790,430
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(19)
OTHER INFORMATION: Zea maize genomic and vector DNA
US-10-790-430-9

Query Match 100.0%; Score 19; DB 18; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19
Db 1 TGTAGCGGCCACGCGTGG 19

RESULT 3

US-09-872-051-7
Sequence 7, Application US/09872051
Patent No. US20020013960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 498
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(498)
OTHER INFORMATION: 1-304 Zea maize genomic DNA
OTHER INFORMATION: 305-349 construct vector DNA
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA
US-09-872-051-7

Query Match 100.0%; Score 19; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19
Db 296 TGTAGCGGCCACGCGTGG 314

RESULT 4

US-10-790-430-7
Sequence 7, Application US/10790430
Publication No. US20040139493A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/10/790,430
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/872,051
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 498
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(498)
OTHER INFORMATION: 1-304 Zea maize genomic DNA
OTHER INFORMATION: 305-349 construct vector DNA
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA
US-10-790-430-7

Query Match 100.0%; Score 19; DB 18; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19
Db 296 TGTAGCGGCCACGCGTGG 314

RESULT 5

US-10-156-761-1684/c
Sequence 1684, Application US/10156761
Publication No. US20030119018A1

```
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ PRIOR FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 1684
/ LENGTH: 1044
/ TYPE: DNA
/ ORGANISM: Streptomyces avermitilis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1044)
/ US-10-156-761-1684

Query Match      84.2%; Score 16; DB 15; Length 1044;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 TAGCGGCCCGCGGTG 18
Db      359 TAGCGGCCCGCGGTG 344

RESULT 6
US-10-645-794-5/c
/ Sequence 5, Application US/10645794
/ Publication No. US20040106194A1
/ GENERAL INFORMATION:
/ APPLICANT: Bett, Andrew J.
/ APPLICANT: Chastain, Michael
/ APPLICANT: Sandig, Volker
/ APPLICANT: Emml, Emilio A.
/ APPLICANT: Shiver, John W.
/ APPLICANT: Casimiro, Danilo R.
/ APPLICANT: Kaslow, David C.
/ APPLICANT: Morsy, Manal
/ TITLE OF INVENTION: METHODS FOR PROPAGATING ADENOVIRUS AND
/ TITLE OF INVENTION: VIRUS PRODUCED THEREBY
/ FILE REFERENCE: 20699X
/ CURRENT APPLICATION NUMBER: US/10/645,794
/ CURRENT FILING DATE: 2003-08-21
/ PRIOR APPLICATION NUMBER: 60/458,825
/ PRIOR FILING DATE: 2003-03-28
/ PRIOR APPLICATION NUMBER: 60/455,312
/ PRIOR FILING DATE: 2003-03-17
/ PRIOR APPLICATION NUMBER: 60/455,234
/ PRIOR FILING DATE: 2003-03-17
/ PRIOR APPLICATION NUMBER: 60/405,182
/ PRIOR FILING DATE: 2002-08-22
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 35167
/ TYPE: DNA
/ ORGANISM: adenovirus serotype 24
/ US-10-645-794-5

Query Match      84.2%; Score 16; DB 18; Length 35167;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGTAGCGGCCCGCGCG 16
Db      12091 TGTAGCGGCCCGCGCG 12076

RESULT 7
US-10-645-883A-1/c
/ Sequence 1, Application US/10645883A
/ Publication No. US2004018555A1
/ GENERAL INFORMATION:
/ APPLICANT: Emml, Emilio A.
/ APPLICANT: Shiver, John W.
/ APPLICANT: Bett, Andrew J.
/ APPLICANT: Casimiro, Danilo R.
/ APPLICANT: Chastain, Michael
/ APPLICANT: Kaslow, David C.
/ APPLICANT: Morsy, Manal
/ TITLE OF INVENTION: ADENOVIRUS SEROTYPE 24 VECTORS, NUCLEIC
/ TITLE OF INVENTION: ACIDS AND VIRUS PRODUCED THEREBY
/ FILE REFERENCE: 21366
/ CURRENT APPLICATION NUMBER: US/10/645,883A
/ CURRENT FILING DATE: 2003-08-21
/ PRIOR APPLICATION NUMBER: 60/455,312
/ PRIOR FILING DATE: 2003-03-17
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 35167
/ TYPE: DNA
/ ORGANISM: adenovirus serotype 24
/ US-10-645-883A-1

Query Match      84.2%; Score 16; DB 18; Length 35167;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGTAGCGGCCCGCGCG 16
Db      12091 TGTAGCGGCCCGCGCG 12076

RESULT 8
US-10-156-761-1
/ Sequence 1, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 1
/ LENGTH: 9025608
/ TYPE: DNA
/ ORGANISM: Streptomyces avermitilis
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (4187715)
/ OTHER INFORMATION: a, t, c, g, other or unknown
/ US-10-156-761-1

Query Match      84.2%; Score 16; DB 15; Length 9025608;
```

Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCCAACGGCTG 18
DB 2073859 TAGCGGCCCAACGGCTG 2073874

RESULT 9

US-10-424-599-43339/c
; Sequence 43339, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43339
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT847_139133C.1
US-10-424-599-43339

Query Match 78.9%; Score 15; DB 17; Length 521;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCCAACGGCT 17
DB 195 TAGCGGCCCAACGGCT 181

RESULT 10
US-10-767-701-9887/c
; Sequence 9887, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9887
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS9666_1
US-10-767-701-9887

Query Match 78.9%; Score 15; DB 18; Length 797;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCCAACGGC 16
DB 195 GTAGCGGCCCAACGGC 181

RESULT 11
US-10-437-963-26588/c

Sequence 26588, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 26588
; LENGTH: 1339
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31363C.1
US-10-437-963-26588

Query Match 78.9%; Score 15; DB 18; Length 1339;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCAACGGCTG 18
DB 121 AGCGGCCCAACGGCTG 107

RESULT 12
US-10-437-963-33098/c
; Sequence 33098, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33098
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31243C.1
US-10-437-963-33098

Query Match 78.9%; Score 15; DB 18; Length 1398;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCAACGGCTG 18
DB 617 AGCGGCCCAACGGCTG 603

RESULT 13
US-10-473-390-1/c
; Sequence 1, Application US/10473390

```
; Publication No. US20040248112A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001206-US
; CURRENT APPLICATION NUMBER: US/10/473,390
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/US02/09743
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,923
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1780
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-390-1
```

```
Query Match      78.9%; Score 15; DB 18; Length 1780;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 AGCGGCCCAAGCGGTG 18
      |||||
Db      51 AGCGGCCCAAGCGGTG 37
```

```
RESULT 14
US-10-437-963-60106
; Sequence 60106, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barzak, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 60106
; LENGTH: 5013
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61666C.1
US-10-437-963-60106
```

```
Query Match      78.9%; Score 15; DB 18; Length 5013;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 GTAGCGGCCCAAGCGG 16
      |||||
Db      2857 GTAGCGGCCCAAGCGG 2871
```

```
RESULT 15
US-10-473-390-3/C
; Sequence 3, Application US/10473390
; Publication No. US20040248112A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
```

```
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001206-US
; CURRENT APPLICATION NUMBER: US/10/473,390
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/US02/09743
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,923
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-390-3
```

```
Query Match      78.9%; Score 15; DB 18; Length 14800;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 AGCGGCCCAAGCGGTG 18
      |||||
Db      2050 AGCGGCCCAAGCGGTG 2036
```

```
Search completed: February 9, 2005, 11:12:40
Job time : 43.5981 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 18.6055 Seconds
(without alignments)
5727.097 Million cell updates/sec

Title: US-10-790-430-10

Sequence: 18
1 taccacgcgcacacttc 18

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	6	ABK15247
2	18	100.0	18	6	ABK15245
3	16	88.9	1338	13	ADSS7504
4	15	83.3	687	3	AAF08489
5	15	83.3	747	4	AAF71736
6	15	83.3	309400	5	AAH68534
7	15	83.3	349980	5	AAH68533
8	14	77.8	320	13	ADQ54932
9	14	77.8	435	5	ABV40711
10	14	77.8	435	5	ABV31743
11	14	77.8	435	5	ABN75645
12	14	77.8	437	5	ABV01404
13	14	77.8	484	5	ABV10573
14	14	77.8	732	6	ABQ14948
15	14	77.8	732	6	ABQ14949
16	14	77.8	2000	8	ADA72018
17	14	77.8	5813	4	ABL07158
18	13	72.2	233	2	AAV20257
19	13	72.2	279	6	ABN16209
20	13	72.2	296	10	ABX85992

ALIGNMENTS

21	13	72.2	314	6	ABQ98333	Abq98333 Human ORF
22	13	72.2	419	13	ADSS4342	Adss4342 Bacterial
23	13	72.2	729	6	ABQ45115	Abq45115 Oligonuc1
24	13	72.2	729	6	ABQ45114	Abq45114 Oligonuc1
25	13	72.2	840	6	ABQ35494	Abq35494 Oligonuc1
26	13	72.2	840	6	ABQ35495	Abq35495 Oligonuc1
27	13	72.2	904	6	ABQ35644	Abq35644 Oligonuc1
28	13	72.2	904	6	ABQ35645	Abq35645 Oligonuc1
29	13	72.2	1344	10	ADG73042	Adg73042 DNA encod
30	13	72.2	1344	10	ADL12097	Adl12097 Pseudomon
31	13	72.2	1396	6	ABQ54082	Abq54082 Oligonuc1
32	13	72.2	1396	6	ABQ54083	Abq54083 Oligonuc1
33	13	72.2	1467	5	AA888878	Aa888878 DNA encod
34	13	72.2	1620	6	ACA26538	Ac26538 Prokaryot
35	13	72.2	2000	6	ABZ17153	Abz17153 Arabidops
36	13	72.2	2761	4	ABL04601	Ab104601 Drosophi1
37	13	72.2	3090	4	ABL15163	Ab115163 Drosophi1
38	13	72.2	4682	4	ABL04600	Ab104600 Drosophi1
39	13	72.2	5580	4	ABL15162	Ab115162 Drosophi1
40	13	72.2	8246	6	ABL32202	Ab132202 Human imm
41	13	72.2	9080	4	ABL19130	Ab119130 Drosophi1
42	13	72.2	13875	4	AA859616	Aa859616 Propionib
43	13	72.2	13875	8	ACF64545	Ac64545 Propionib
44	13	72.2	17369	4	ABL02754	Ab102754 Drosophi1
45	13	72.2	28199	8	ADA68658	Ada68658 Takifugu

RESULT 1	ABK15247	standard, DNA, 18 BP.
ID	ABK15247	
XX	ABK15247	
AC	ABK15247	
DT	08-MAY-2002	(first entry)
XX	XX	
DE	Corn transgene junction probe #2.	
XX	XX	
KW	Probe; ss; rice actin 1 promoter; RA1; RA1 intron;	
KW	chloroplast transit peptide gene; glyphosate resistance; corn;	
KW	5-enol-pyruvylphosphoketide-3-phosphatase synthase; EPSPS; Hsp70 intron;	
KW	transcriptional terminator; cauliflower mosaic virus 35S promoter;	
KW	PV-ZWGT32; transgenic; nk603.	
XX	XX	
OS	Zea mays.	
OS	Synthetic.	
XX	XX	
PN	EP167531-A1.	
XX	XX	
PD	02-JAN-2002.	
XX	XX	
PF	15-JUN-2001; 2001EP-00202314.	
XX	XX	
PR	22-JUN-2000; 2000US-0213567P.	
PR	13-OCT-2000; 2000US-0240014P.	
PR	13-OCT-2000; 2000US-0241215P.	
XX	XX	
PA	(MONS) MONSANTO TECHNOLOGY LLC.	
XX	XX	
PI	Behr CF, Hitonaka C, Heck GR, You J;	
XX	XX	
DR	WPI; 2002-165871/22.	
XX	XX	
PT	Novel DNA construct useful for producing a corn plant that tolerates	
PT	application of glyphosate herbicide, comprises two transgene expression	
PT	cassettes.	
XX	XX	
PS	Claim 8; Page 3; 25pp; English.	
XX	XX	
CC	The invention relates to a DNA construct (1) comprising a first (F1) and	
CC	a second (F2) expression cassette. F1 of the DNA construct, in operable	

Query Match	Best Local Similarity	100.0%	Score 18;	DB 6;	Length 18;
Matches	18;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0
Qy	1	TACCAAGCAGCACACTTC	18		
Db	1	TACCAAGCAGCACACTTC	18		
RESULT 2					
ABK15245					
ID	ABK15245	standard; DNA; 1183 BP.			
AC	ABK15245;				
DT	08-MAY-2002	(first entry)			
XX					
DE	Corn nk603 3' transgene/genomic sequence PCR fragment.				
XX					
KW	de; rice actin 1 promoter; RAI; RAI intron;				
KW	chloroplast transit peptide gene; glyphosate resistance; corn;				
KW	5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; HEP70 intron;				
KW	transcriptional terminator; cauliflower mosaic virus 35S promoter;				
KW	PV-2MG332; transgenic; nk603.				
XX					
OS	Zea mays.				
OS	Escherichia coli.				
OS	Agrobacterium tumefaciens.				
OS	Synthetic.				
OS	Chimeric.				
XX					
PH	Key	Location/Qualifiers			
FT	misc_feature	1..164			
FT		/*tag= a			
FT		/note= "T-AGRTU.nos vector sequence"			
FT	misc_feature	165..381			
FT		/*tag= b			
FT		/note= "Vector sequence"			
FT	misc_feature	382..686			
FT		/*tag= c			
FT		/note= "Corn plastid genes rps11/rpoA partial sequence"			
FT	misc_feature	687..1183			
FT		/*tag= d			
FT		/note= "Corn genomic sequence"			
XX					
PN	EP1167531-A1.				
XX					
PD	02-JAN-2002.				
XX					
XX	15-JUN-2001; 2001EP-00202314.				
XX					
PR	22-JUN-2000; 2000US-0213567P.				
PR	13-OCT-2000; 2000US-0240014P.				
PR	13-OCT-2000; 2000US-0241215P.				
XX					
PA	(MONS) MONSANTO TECHNOLOGY LLC.				
XX					
XX	Behr CF, Hironaka C, Heck GR, You J;				

DR	XX	WP1; 2002-165871/22.
PT	XX	Novel DNA construct useful for producing a corn plant that tolerates application of glyphosate herbicide, comprises two transgene expression cassettes.
PS	XX	Claim 3; Page 16-17; 25pp; English.
CC	XX	The invention relates to a DNA construct (I) comprising a first (F1) and a second (F2) expression cassette. F1 of the DNA construct, in operable linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-phosphate synthase (EPSPS) gene (III), and transcriptional terminator (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron, (II), (III) and (IV). Also included are two DNA sequences from corn plant PV-ZMOT32(nk603) of 498 or 1183bp, or primers and probes derived from them used to detect transgene junction points). The construct is useful for producing a corn plant that tolerates application of glyphosate herbicide by transforming a corn cell with, selecting the corn cell for tolerance to application of glyphosate, and growing the corn cell into a fertile corn plant. The present sequence is the 1183bp amplicon consisting of the Agrobacterium transcriptional terminator sequence, vector sequences and corn sequences
SO	XX	Sequence 1183 BP; 295 A; 289 C; 280 G; 319 T; 0 U; 0 Other;
QY	Db	Query Match 100.0%; Score 18; DB 6; Length 1183; Best Local Similarity 100.0%; Pred. No. 0.17; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
		1 TACCACGCGACACTTC 18 373 TACCACGCGACACTTC 390
RESULT 3		
AD557504/c		
ID	AD557504	standard; cDNA; 1338 BP.
XX	AC	
XX	AD557504;	
XX	DT	02-DEC-2004 (first entry)
DE	XX	Bacterial polynucleotide #9491.
XX	XX	Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.
OS	XX	Bacteria.
XX	PN	US2003233675-A1.
XX	PD	18-DEC-2003.
XX	PF	20-FEB-2003; 2003US-00369493.
XX	PR	21-FEB-2002; 2002US-0360039P.
XX	PA	(CAOY/) CAO Y.
XX	PA	(HINK/) HINKLE G J.
XX	PA	(SLAT/) SLATER S C.
XX	PA	(CHEN/) CHEN X.
XX	PA	(GOLD/) GOLDMAN B S.
XX	PI	Gao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	OR	WP1; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
PS Claim 1, SEQ ID NO 33178; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plants with
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomanan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/Sequence.html.
XX
SQ Sequence 1338 BP; 241 A; 281 C; 463 G; 353 T; 0 U; 0 Other;
XX
Query Match 88.9%; Score 16; DB 13; Length 1338;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2 ACCACGGCAGACACTT 17
Db 513 ACCACGGCAGACACTT 498
XX
RESULT 4
AAF08489
ID AAF08489 standard; cDNA; 687 BP.
XX
AC AAF08489;
XX
DT 13-MAR-2001 (first entry)
XX
DE Fusarium venenatum EST SEQ ID NO:1012.
XX
XX Multiple gene expression; filamentous fungal cell, EST;
KW expressed sequence tag; Fusarium venenatum; *Aspergillus niger*;
KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
XX *Fusarium venenatum*.
OS
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US007781.
XX
PR 22-MAR-1999; 99US-00273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

XX
DR MPI; 2000-594572/56.
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags.
XX
XX Claim 86; Page 768; 3161pp; English.
PS
XX The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring the
CC global expression of genes from FF cells allows the production potential
CC of the microorganisms to be improved. New genes may be discovered,
CC possible functions of unknown open reading frames can be identified and
CC gene copy number variation and stability can be monitored. The expression
CC of genes can be used to study how FF cells adapt to changes in culture
CC conditions, environmental stress, spore morphogenesis, recombination,
CC metabolic or catabolic pathway engineering. Using ESTs provides several
CC advantages over genomic or random cDNA clones including elimination of
CC redundancy as one spot on an array equals one gene or open reading frame,
CC and organization of the microarrays based on function of the gene
CC products to facilitate analysis of the results. AAF07478 to AAF11247
CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents
CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from
CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from
CC *Trichoderma reesei*, which are all specifically claimed in the present
CC invention
XX
SQ Sequence 687 BP; 167 A; 190 C; 152 G; 178 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 3; Length 687;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 4 CACGCGACACACTTC 18
Db 531 CACGCGACACACTTC 545
XX
RESULT 5
AAF71736/C
ID AAF71736 standard; DNA; 747 BP.
XX
AC AAF71736;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:753.
XX
XX Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
KW diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
XX *Corynebacterium glutamicum*.
OS
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000943.
XX
PR 25-JUN-1999; 99US-0141031P.
XX
PR 08-JUN-1999; 99DE-01031412.

PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031564.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX (BADT) BASF AG.
XX
XX
PI Pompeius M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-061975/07.
XX P-PSDB; AAB79619.
XX
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes.
XX
XX
PS Claim 3; Page 1209-1210; 1246pp; English.
XX
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for
CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteihogenic or
CC nonproteihogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum sequences of interest, in
CC evolutionary studies, in determining SMP protein sequences required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH)
XX
XX
SQ Sequence 747 BP; 175 A; 169 C; 205 G; 198 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 4; Length 747;
Best Local Similarity 100.0%; Pred. No. 12;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACCGACACACTTC 18
|||
DB 21 CACCGACACACTTC 7

RESULT 6
AAH68534/C
ID AAH68534 standard; DNA; 309400 BP.
XX
XX
AC AAH68534;
XX
DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7069.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis; de.

OS Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tareishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.

XX Disclosure; SEQ ID NO 7069; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 5; Length 309400;
Best Local Similarity 100.0%; Pred. No. 11;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACCGACACACTTC 18
|||
DB 28188 CACCGACACACTTC 28174

RESULT 7

AAH68533/C
ID AAH68533 standard; DNA; 349980 BP.

XX AAH68533;
XX

DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7068.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.
 XX Corynebacterium glutamicum.
 XX EP1108790-A2.
 XX 20-JUN-2001.
 XX 18-DEC-2000; 2000EP-00127688.
 XX 16-DEC-1999; 99JP-00377484.
 XX 07-APR-2000; 2000JP-00159162.
 XX 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 DR WPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Disclosure; SEQ ID NO 7068; 246bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;
 Query Match 83.3%; Score 15; DB 5; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CACGCGACACACTTC 18
 DB 328188 CACGCGACACACTTC 328174
 RESULT 8
 ADQ54932/C
 ID ADQ54932 standard; DNA; 320 BP.
 XX
 AC ADQ54932;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DB Novel canine microarray-related DNA sequence SeqID6234.
 XX
 KW canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; ds.
 XX
 OS Canis familiaris.
 XX
 PN WO2004063324-A2.
 XX
 PD 29-JUL-2004.
 XX
 PP 05-MAY-2003; 2003WO-US013853.

XX
 PR 03-MAY-2002; 2002US-0377240P.
 XX
 PA (GENE-) GENE LOGIC INC.
 PA (PF12) PFIZER PROD INC.
 XX
 PI Digians JC, Porter M, Wei T;
 XX
 DR WPI; 2004-561890/54.
 XX
 PT New isolated nucleic acid molecule, useful for drug screening and
 PT toxicity assays or for assessing the impact, including toxicity, of a
 PT compound, pharmaceutical agent or environmental pollutant on a cell or
 PT living organism.
 XX
 PS Claim 1; SEQ ID NO 6234; 41pp; English.
 XX
 CC This invention is related to a novel isolated canine nucleic acid
 CC sequences and the construction of canine microarrays containing a
 CC significant portion of the canine genome. The isolated canine nucleic
 CC acid sequences of the invention may be useful for drug screening and
 CC toxicity assays. The invention is therefore useful for assessing the
 CC impact, including toxicity, of a compound, pharmaceutical agent or
 CC environmental pollutant on a cell or living organism. The methods are
 CC useful for detecting genes that are up- or down-regulated in canines in a
 CC disease state. The sequences are useful as diagnostic agents or markers
 CC to detect a cellular response in a sample individually or as part of a
 CC gene expression profile. It is also useful as a target for agents that
 CC modulate gene expression or activity. The database is useful for
 CC producing electronic Northern blots that allow the user to determine the cell
 CC type or tissue in which a given gene is expressed and to allow
 CC determination of the abundance or expression level of a given gene in a
 CC particular tissue or cell. The methods are useful for determining the
 CC similarity of a toxic response to one or more individual compounds. The
 CC methods are useful for predicting at least one toxic response or the
 CC likelihood that a compound or test agent will induce various specific
 CC pathologies such as those of the liver (liver necrosis, fatty liver
 CC disease, protein adduct formation or hepatitis), those of the kidney,
 CC heart, brain or testes, or other pathologies associated with at least one
 CC of the toxins. The methods are also useful for predicting or elucidating
 CC the potential cellular pathways influenced, induced or modulated by the
 CC compound or test agent due to the similarity of the expression profile
 CC compared to the profile induced by a known toxin. The present sequence is
 CC that of a canine DNA sequence which was claimed for use during the
 CC production of a canine microarray of the invention.
 XX
 SQ Sequence 320 BP; 35 A; 97 C; 134 G; 49 T; 0 U; 5 Other;
 Query Match 77.8%; Score 14; DB 13; Length 320;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ACCACGCGACACAC 15
 DB 230 ACCACGCGACACAC 217
 RESULT 9
 ABV40711/C
 ID ABV40711 standard; cDNA; 435 BP.
 XX
 AC ABV40711;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 40702.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.

XX 23-AUG-2001.
PD 20-FEB-2001; 2001WO-US005171.
XX
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8202; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 435 BP; 128 A; 102 C; 110 G; 95 T; 0 U; 0 Other;
XX
Query Match 77.8%; Score 14; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CACGCGACACACTT 17
DB 286 CACGCGACACACTT 273
XX
RESULT 10
ABV31743/c
ID ABV31743 standard; cDNA; 435 BP.
XX
AC ABV31743;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 31734.
XX
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 6811; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 435 BP; 128 A; 102 C; 110 G; 95 T; 0 U; 0 Other;
XX
Query Match 77.8%; Score 14; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CACGCGACACACTT 17
DB 286 CACGCGACACACTT 273
XX
RESULT 11
ABN75645
ID ABN75645 standard; cDNA; 435 BP.
XX
AC ABN75645;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human ORF592 cDNA, SEQ ID NO:1183.
XX
XX Human; ORF; open reading frame; ORF; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulvectomy;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; anticholesterol; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017076.
XX
PR 24-MAY-2000; 2000US-0206690P.
XX

PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR
XX
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer; useful
PT for detecting presence of prostate cancer; stage of prostate cancer.
XX
PS Claim 1; Page 1705; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 484 BP; 140 A; 113 C; 124 G; 107 T; 0 U; 0 Other;
XX
Query Match 77.8%; Score 14; DB 5; Length 484;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CACGCGACACACTT 17
DB 335 CACGCGACACACTT 322
RESULT 14
ABQ14948/c
ID ABQ14948 standard; DNA; 732 BP.
XX
AC ABQ14948;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 1539.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.

XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 732 BP; 64 A; 82 C; 341 G; 245 T; 0 U; 0 Other;
XX
Query Match 77.8%; Score 14; DB 6; Length 732;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ACCACGCGACACAC 15
DB 548 ACCACGCGACACAC 535
RESULT 15
ABQ14949
ID ABQ14949 standard; DNA; 732 BP.
XX
AC ABQ14949;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 1540.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons

PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX
PS
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABO54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention

XX Sequence 732 BP; 245 A; 341 C; 82 G; 64 T; 0 U; 0 Other;

XX
SQ Query Match 77.8%; Score 14; DB 6; Length 732;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACCACGGGACACAC 15
|||||

Db 185 ACCACGGGACACAC 198

Search completed: February 10, 2005, 13:44:08

Job time : 22.6055 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 29.935 Seconds
(without alignments)
3460.943 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18

Sequence: 1 tgcgtctcgtcgtacttc 18

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 4313806 seqs, 287787103 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	9	US-09-872-051-11
2	18	100.0	18	18	US-10-790-430-11
3	18	100.0	1183	9	US-09-872-051-8
4	18	100.0	1183	18	US-10-790-430-8
5	18	100.0	1183	18	US-10-425-115-84576
6	15	83.3	21	17	US-10-349-143-7605
7	15	83.3	21	17	US-10-357-930-19880
8	15	83.3	445	18	US-10-357-930-49671
9	15	83.3	500	18	US-10-027-632-234880
10	15	83.3	533	17	US-10-027-632-234880
11	15	83.3	641	18	US-10-767-701-6161

12	15	83.3	695	17	US-10-424-599-77613	Sequence 77613, A
13	15	83.3	3167	10	US-09-764-891-8948	Sequence 8948, Ap
14	15	83.3	70019	18	US-10-323-281-823	Sequence 823, App
15	15	77.8	20	17	US-10-289-752-5457	Sequence 5457, Ap
16	15	77.8	25	18	US-10-717-597-3196	Sequence 3196, Ap
17	15	77.8	25	18	US-10-717-597-3197	Sequence 3197, Ap
18	15	77.8	149	18	US-10-357-930-15466	Sequence 15466, A
19	15	77.8	154	18	US-10-357-930-15388	Sequence 15388, A
20	15	77.8	169	9	US-09-864-761-26676	Sequence 26676, A
21	14	77.8	201	18	US-10-741-601-5342	Sequence 5342, Ap
22	14	77.8	201	18	US-10-741-601-5350	Sequence 5350, Ap
23	14	77.8	201	18	US-10-741-601-20494	Sequence 20494, A
24	14	77.8	230	18	US-10-357-930-36207	Sequence 36207, A
25	14	77.8	241	17	US-10-305-720-403	Sequence 403, App
26	14	77.8	257	18	US-10-674-124A-21610	Sequence 21610, A
27	14	77.8	366	17	US-10-282-122A-31878	Sequence 31878, A
28	14	77.8	379	18	US-10-357-930-36285	Sequence 36285, A
29	14	77.8	400	10	US-09-918-995-33608	Sequence 33608, A
30	14	77.8	400	17	US-10-242-535A-20475	Sequence 20475, A
31	14	77.8	400	17	US-10-085-783A-20475	Sequence 20475, A
32	14	77.8	426	17	US-10-282-122A-31877	Sequence 31877, A
33	14	77.8	435	17	US-10-242-535A-38277	Sequence 38277, A
34	14	77.8	435	17	US-10-085-783A-38277	Sequence 38277, A
35	14	77.8	441	17	US-10-242-535A-25043	Sequence 25043, A
36	14	77.8	441	17	US-10-085-783A-25043	Sequence 25043, A
37	14	77.8	450	17	US-10-424-599-110092	Sequence 110092, A
38	14	77.8	483	16	US-10-029-366-23957	Sequence 23957, A
39	14	77.8	496	9	US-09-864-761-10034	Sequence 10034, A
40	14	77.8	496	18	US-10-357-930-45260	Sequence 45260, A
41	14	77.8	496	18	US-10-357-930-45326	Sequence 45326, A
42	14	77.8	515	11	US-09-864-408A-8531	Sequence 8531, Ap
43	14	77.8	547	16	US-10-029-366-10240	Sequence 10240, A
44	14	77.8	549	10	US-09-991-935-126	Sequence 126, App
45	14	77.8	573	18	US-10-021-323-14061	Sequence 14061, A

ALIGNMENTS

RESULT 1

US-09-872-051-11

Sequence 11, Application US/09872051

Patent No. US20020013960A1

GENERAL INFORMATION:

APPLICANT: Monsanto Co

APPLICANT: Behr, Carl

APPLICANT: Hironaka, Catherine

APPLICANT: Heck, Gregory

APPLICANT: You, Jinsong

TITLE OF INVENTION: Corn Event PV-ZMGR32 (n603) and Composition and Methods for Det

FILE REFERENCE: 38-21(52258)B

CURRENT APPLICATION NUMBER: US/09/872,051

PRIOR FILING DATE: 2001-06-01

PRIOR APPLICATION NUMBER: 60/213,567

PRIOR FILING DATE: 2000-06-22

PRIOR APPLICATION NUMBER: 60/241,215

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: 60/240,014

PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.0

SEQ ID NO 11

LENGTH: 18

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: source

LOCATION: (1)..(18)

OTHER INFORMATION: Zea maize genomic DNA and vector DNA

US-09-872-051-11

Query Match

100.0%;

Score 18;

DB 9;

Length 18;

Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
| | | | | | | | | | | | | | | | | |
Db 1 TGCTGTTCTGCTGACTTT 18

RESULT 2

US-10-790-430-11
; Sequence 11, Application US/10790430
; Publication No. US20040139493A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/10/790,430
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(18)
; OTHER INFORMATION: Zea maize genomic DNA and vector DNA
US-10-790-430-11

Query Match 100.0%; Score 18; DB 18; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
| | | | | | | | | | | | | | | | | |
Db 1 TGCTGTTCTGCTGACTTT 18

RESULT 3

US-09-872-051-8
; Sequence 8, Application US/09872051
; Patent No. US20020013960A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plasmid genes, rps11 and rpsA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

Query Match 100.0%; Score 18; DB 9; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
| | | | | | | | | | | | | | | | | |
Db 678 TGCTGTTCTGCTGACTTT 695

RESULT 4

US-10-790-430-8
; Sequence 8, Application US/10790430
; Publication No. US20040139493A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/10/790,430
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plasmid genes, rps11 and rpsA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-10-790-430-8

Query Match 100.0%; Score 18; DB 18; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18
| | | | | | | | | | | | | | | | | |
Db 678 TGCTGTTCTGCTGACTTT 695

RESULT 5

US-10-425-115-84576

```
; Sequence 84576, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 84576
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1818)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_177144C.1
; US-10-425-115-84576
```

```
Query Match      88.9%; Score 16; DB 18; Length 1818;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGCTGTTCTGCTGACT 16
          |||||
DB      59 TGCTGTTCTGCTGACT 74
```

```
RESULT 6
; US-10-349-143-7605
; Sequence 7605, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CP1
; CURRENT APPLICATION NUMBER: US/10/349,143
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US/09/422,978
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 7605
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1..21
; OTHER INFORMATION: upstream amplification primer 99-9623 for SEQ 3671,
; US-10-349-143-7605
```

```
Query Match      83.3%; Score 15; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CTGTTCTGCTGACTT 17
          |||||
DB      2 CTGTTCTGCTGACTT 16
```

```
RESULT 7
; US-10-357-930-19880/C
; Sequence 19880, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19880
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-357-930-19880
```

```
Query Match      83.3%; Score 15; DB 18; Length 445;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGCTGTTCTGCTGAC 15
          |||||
DB      256 TGCTGTTCTGCTGAC 242
```

```
RESULT 8
; US-10-357-930-49671/C
; Sequence 49671, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
```

;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 49671
;; LENGTH: 500
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-357-930-49671

Query Match 83.3%; Score 15; DB 18; Length 500;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGAC 15
Db 296 TGCTGTTCTGCTGAC 282

RESULT 9
US-10-027-632-234880/c
;; Sequence 234880, Application US/10027632
;; Publication No. US20020198371A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 234880
;; LENGTH: 533
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-234880

Query Match 83.3%; Score 15; DB 13; Length 533;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TGTTCTGCTGACTT 18
Db 374 TGTTCTGCTGACTT 360

RESULT 10
US-10-027-632-234880/c
;; Sequence 234880, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12

;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 234880
;; LENGTH: 533
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-234880

Query Match 83.3%; Score 15; DB 17; Length 533;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TGTTCTGCTGACTT 18
Db 374 TGTTCTGCTGACTT 360

RESULT 11
US-10-767-701-6161
;; Sequence 6161, Application US/10767701
;; Publication No. US20040172684A1
;; GENERAL INFORMATION:
;; APPLICANT: Kovacic, David K.
;; APPLICANT: Zhou, Yihua
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;; FILE REFERENCE: 38-21(53535)B
;; CURRENT APPLICATION NUMBER: US/10/767,701
;; CURRENT FILING DATE: 2004-01-29
;; NUMBER OF SEQ ID NOS: 63128
;; SEQ ID NO 6161
;; LENGTH: 641
;; TYPE: DNA
;; ORGANISM: Sorghum bicolor
;; FEATURE:
;; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS35548_1
US-10-767-701-6161

Query Match 83.3%; Score 15; DB 18; Length 641;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGAC 15
Db 586 TGCTGTTCTGCTGAC 600

RESULT 12
US-10-424-599-77613
;; Sequence 77613, Application US/10424599
;; Publication No. US20040031072A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa Thomas J
;; APPLICANT: Kovalic David K
;; APPLICANT: Zhou Yihua
;; APPLICANT: Cao Yongwei
;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;; FILE REFERENCE: 38-21(53223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599

```
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 77613
/ LENGTH: 695
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_41100C.1
US-10-424-599-77613

Query Match      83.3%; Score 15; DB 17; Length 695;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGCTGTTCTGCTGAC 15
         |||||
Db      187 TGCTGTTCTGCTGAC 201

RESULT 13
US-09-764-891-8948
/ Sequence 8948, Application US/09764891
/ Publication No. US20030077808A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC006
/ CURRENT APPLICATION NUMBER: US/09/764,891
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 10231
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 8948
/ LENGTH: 3167
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-764-891-8948

Query Match      83.3%; Score 15; DB 10; Length 3167;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 TGTTCTGCTGACTTT 18
         |||||
Db      1739 TGTTCTGCTGACTTT 1753

RESULT 14
US-10-322-281-823
/ Sequence 823, Application US/10322281
/ Publication No. US20040126762A1
/ GENERAL INFORMATION:
/ APPLICANT: David W. Morris
/ APPLICANT: Marc S. Malandro
/ TITLE OF INVENTION: Novel Compositions and Methods in Cancer
/ FILE REFERENCE: 529452001000
/ CURRENT APPLICATION NUMBER: US/10/322,281
/ CURRENT FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 823
/ LENGTH: 70019
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-10-322-281-823

Query Match      83.3%; Score 15; DB 18; Length 70019;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 TGTTCTGCTGACTTT 18
         |||||
Db      27137 TGTTCTGCTGACTTT 27151
```

```
RESULT 15
US-10-289-762-5457/C
/ Sequence 5457, Application US/10289762
/ Publication No. US20040006218A1
/ GENERAL INFORMATION:
/ APPLICANT: Grifais, R
/ TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragmen-
/ thereof and uses thereof, in particular for the diagnosis, pre-
/ FILE REFERENCE: 9710-003-999
/ CURRENT APPLICATION NUMBER: US/10/289,762
/ CURRENT FILING DATE: 2003-03-27
/ NUMBER OF SEQ ID NOS: 6849
/ SEQ ID NO 5457
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Chlamydia pneumoniae
US-10-289-762-5457

Query Match      77.8%; Score 14; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 GTTCTGCTGACTTT 18
         |||||
Db      16 GTTCTGCTGACTTT 3
```

Search completed: February 9, 2005, 11:12:46
Job time : 30.935 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 111.027 Seconds
(without alignments)
7855.666 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18

Sequence: 1 accaagctttataatag 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:*
1: gb_Da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	6	AX342373 Sequence
2	18	100.0	1183	6	AX342369 Sequence
3	18	100.0	225438	2	AC106676 Rattus no
4	18	100.0	232994	2	AC110467 Rattus no
5	17	94.4	118174	2	AC138014 Medicago
6	17	94.4	158959	9	AC004925 Homo sapi
7	17	94.4	175940	9	AL133551 Human DNA
8	16	88.9	42563	1	AY160838 Uncultured
9	16	88.9	48843	8	AB018110 Arabidops
10	16	88.9	48843	8	AF467900 Prunus pe
11	16	88.9	51797	2	AC145628 Homo sapi
12	16	88.9	58068	2	AC109350 Homo sapi
13	16	88.9	107044	9	AC092284 Homo sapi
14	16	88.9	107129	2	AC092359 Homo sapi
15	16	88.9	133244	9	AC104793 Homo sapi
16	16	88.9	135270	2	AC068366 Homo sapi
17	16	88.9	138215	2	AC023523 Homo sapi
18	16	88.9	141645	2	AC125804 Mus muscu
19	16	88.9	143575	2	AC092370 Homo sapi

c	20	16	88.9	150102	2	BX957301 Danio rer
c	21	16	88.9	151699	10	AC110896 Mus muscu
c	22	16	88.9	154622	2	AC136599 Homo sapi
c	23	16	88.9	155611	9	AC007358 Homo sapi
c	24	16	88.9	156764	2	CR392043 Danio rer
c	25	16	88.9	158951	2	CR318594 Danio rer
c	26	16	88.9	165178	2	AC127080 Rattus no
c	27	16	88.9	166688	2	CR382375 Danio rer
c	28	16	88.9	169633	9	AC023235 Homo sapi
c	29	16	88.9	173395	2	AC093406 Mus muscu
c	30	16	88.9	176422	2	AC019138 Homo sapi
c	31	16	88.9	186648	9	AC113407 Homo sapi
c	32	16	88.9	186797	2	AC021965 Homo sapi
c	33	16	88.9	197027	2	AC148230 Gallus ga
c	34	16	88.9	198146	9	AC094019 Homo sapi
c	35	16	88.9	212720	2	AC126574 Rattus no
c	36	16	88.9	212895	2	AC148788 Oryzomur
c	37	16	88.9	213012	2	AC126294 Rattus no
c	38	16	88.9	215731	2	AC136554 Rattus no
c	39	16	88.9	226839	2	AC150579 Bos tauru
c	40	16	88.9	227182	5	BX005355 Zebrafish
c	41	16	88.9	282292	2	AC095274 human STS W
c	42	15	83.3	187	11	G06292
c	43	15	83.3	549	6	AX435036 Sequence
c	44	15	83.3	569	1	233031 M.capricolu
c	45	15	83.3	1177	8	MDAJ761 Matus dom

ALIGNMENTS

RESULT 1	AX342373	18 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342373	Sequence 12 from Patent EP167531.			
DEFINITION	AX342373	AX342373.1 GI:18151816			
ACCESSION	AX342373				
VERSION	AX342373.1	GI:18151816			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE	1				
AUTHORS	Behr, C.F., Hironaka, C., Heck, G.R. and You, J.				
TITLE	Corn transformant pv-zmgf32 (ink03) and compositions and methods for detection thereof				
JOURNAL	Patent: EP 1167531-A 12 02-JAN-2002;				
FEATURES	Monsanto Technology LLC (US)				
LOCATION	Location/Qualifiers				
SOURCE	1. 18				
ORIGIN	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="Agrobacterium tumefaciens nos 3' terminator and rice actin promot er DN"				
Query Match	100.0%;	Score 18;	DB 6;	Length 18;	
Best local Similarity	100.0%;	Pred. No. 16;			
Matches	18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	ACCAAGCTTTATATAG 18			
Db	1	ACCAAGCTTTATATAG 18			
RESULT 2	AX342369	1183 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342369	Sequence 8 from Patent EP167531.			
DEFINITION	AX342369	AX342369.1 GI:18151812			
ACCESSION	AX342369				
VERSION	AX342369.1	GI:18151812			
KEYWORDS					
SOURCE					
ORIGIN					
	synthetic construct				

ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1 Behr, C.F., Hironaka, C., Heck, G.R. and You, J.
Corn transformant pv-rc32 (HK603) and compositions and methods
for detection thereof
Patent: EP 1167531-A 8 02-JAN-2002;
Monsanto Technology LLC (US)

JOURNAL
Location/Qualifiers

FEATURES
Source
1. 1183
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="1-164 Agrobacterium tumefaciens nos 3' terminator
165-381 construct vector DNA 382-686 Zea mays genomic DNA"
genes, rp11 and rpo4 687-1183 Zea mays genomic DNA"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1183;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 ACCAAGCTTTATATAG 18
|||||
156 ACCAAGCTTTATATAG 173

RESULT 3
AC106676 225438 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-107N12, WORKING DRAFT SEQUENCE, 3
DEFINITION unorderd pieces.
AC106676
AC106676 4 GI:30578541
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 225438)
Muzny, D., Maric, M., Mettler, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryan, N., Buhay, C., Burck, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devilla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gubregovis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guerrero, W.,
Guzarane, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangun, A.,
Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Mlooesjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nat, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okwuonu, G., Olarunuga, A., Pal, S., Parke, K.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Paeternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poldexeter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, M., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorreller, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 225438)
Worley, K.C.

Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225438)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23268058.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
table sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLU1
Center clone name: CH230-107N12
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 201968 bases at least Q40
Consensus quality: 205320 bases at least Q30
Consensus quality: 207760 bases at least Q20
Estimated insert size: 213439; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
* 1 222660: contig of 222660 bp in length
* 222661 222760: gap of unknown length
* 222761 224015: contig of 1255 bp in length

FEATURES * 224016 224115: gap of unknown length
 * 224116 225438: contig of 1323 bp in length
 Location/Qualifiers
 source
 1. 225438
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-107N12"
 1. 1186
 /note="wgs end extension
 clone end:Sp6"
 4773_5453
 /note="clone boundary
 clone end:Sp6"
 site:ECORI
 end sequence: BH271428"
 complement(22111..221919)
 /note="clone boundary
 clone end:T7
 site:ECORI
 end sequence: BH271425"

misc_feature
 misc_feature
 misc_feature

ORIGIN
 Query Match 100.0%; Score 18; DB 2; Length 225438;
 Best Local Similarity 100.0%; Pred. No. 5.3; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ACCNAGCTTTATATAG 18
 |||||
 Db 13992 ACCNAGCTTTATATAG 14009

RESULT 4
 AC110467 232994 bp DNA linear HTG 19-SEP-2002
 LOCUS Rattus norvegicus clone CH230-320D17, *** SEQUENCING IN PROGRESS
 DEFINITION *** 7 unordered pieces.
 ACCESSION AC110467
 VERSION AC110467.4 GI:23195262
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 232994)
 Muzny,D.,Marle,M.,Mezker,M.,Lee,A.,Abramson,S.,Adams,C.,Alder,J.,
 Allen,C.,Allen,H.,Albrooke,S.,Amin,A.,Angiano,D.,
 Anyaldechit,V.,Aoyagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
 Baldwin,D.,Bandaraike,D.,Barber,M.,Barnstead,M.,Behamed,F.,
 Blawie,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
 Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderon,E.,
 Cardenas,V.,Carter,K.,Cavazos,I.,Cesar,H.,Center,A.,
 Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
 Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
 Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dedrich,D.,
 Delgado,O.,Denson,S.,Derramo,C.,Ding,Y.,Dinh,H.,Divya,K.,
 Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duvall,B.,Eaves,K.,
 Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
 Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,
 Fraser,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garner,T.,Gara,M.,
 Geisregeorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,
 Gunaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,
 Harvey,Y.,Havlik,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
 Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
 Hollins,B.,Howells,S.,Hulik,S.,Hume,J.,Idlebird,D.,Jackson,A.,
 Jackson,L.,Jacod,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,
 Kapachy,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kovari,C.,
 Kowals,C.,Krafft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,
 Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,
 Lorenshewa,L.,Louisaged,H.,Lozato,R.J.,Lu,X.,Ma,J.,
 Maheshwari,M.,Mahindartne,M.,Mahmond,M.,Malloy,K.,Mangum,A.,
 Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,

REFERENCE
 TITLE
 JOURNAL
 AUTHORS
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
 Nwakoeleneh,O., Okwuonu,G., Olamunsgoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Plankoch,C.,
 Plapper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.,L.,
 Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
 Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Sma3,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steidle,M., Strong,R., Sutton,A., Swatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,K., White,F.,
 Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G., and Gibbs,R.A.
 Direct Submission
 2 (bases 1 to 232994)
 Unpublished
 Morley,K.C.
 Direct Submission
 Submitted (13-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 232994)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 19, 2002 this sequence version replaced gi:21741421.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/atl/). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRM
 Center clone name: CH230-320D17
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 217650 bases at least Q40
 Consensus quality: 220144 bases at least Q30
 Consensus quality: 221530 bases at least Q20
 Bistimulated insert size: 235091; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Bistimulated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdate.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 152446: contig of 152446 bp in length
 * 152447 152546: gap of unknown length
 * 152547 214193: contig of 61647 bp in length

214194 214293: gap of unknown length
* 214294 215380: contig of 1087 bp in length
* 215381 215480: gap of unknown length
* 215481 217850: contig of 2370 bp in length
* 217851 217950: gap of unknown length
* 217951 218965: contig of 1015 bp in length
* 218966 219065: gap of unknown length
* 219066 220974: contig of 1909 bp in length
* 220975 221074: gap of unknown length
* 221075 223294: contig of 11920 bp in length.
Location/Qualifiers
1. .232994
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-320D17"
5849. .7480
/note="wgs_contig"

misc_feature

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 232994;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATATAG 18
|||||
Db 83203 ACCAAGCTTTATATAG 83220

RESULT 5
AC138014 118174 bp DNA linear HTG 04-SEP-2004
LOCUS Medicago truncatula clone mth2-17p13, WORKING DRAFT SEQUENCE, 3
DEFINITION Ordered pieces.
AC138014
AC138014.30 GI:51889812
HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
1 (bases 1 to 118174)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B. A.
Medicago truncatula BAC Clone mth2-17p13
Unpublished
2 (bases 1 to 118174)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B. A.
Direct Submission
Submitted (10-DEC-2002) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
3 (bases 1 to 118174)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B. A.
Direct Submission
Submitted (04-SEP-2004) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
On Sep 4, 2004 this sequence version replaced gi:51699588.
Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code:UOKNOR

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 78995: contig of 78995 bp in length
* 78996 79095: gap of unknown length
* 79096 84864: contig of 5769 bp in length
* 84865 84964: gap of unknown length
* 84965 118174: contig of 33210 bp in length.
Location/Qualifiers
1. .118174
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-17p13"
/clone_11b="Medicago truncatula BAC library H2"

ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 118174;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATATATA 17
|||||
Db 70329 ACCAAGCTTTATATATA 70345

RESULT 6
AC004925 154959 bp DNA linear PRI 30-JUN-2004
LOCUS Homo sapiens PAC clone RP5-907C10 from 7, complete sequence.
AC004925
AC004925.1 GI:4156174
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 154959)
Hillier, L.W., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H., Wagner, M., McKernan, C., Layman, D., Maas, J., Jaeger, S., Walker, R., Wyllie, K., Sekhon, M., Becker, M.C., O'Laughlin, M.D., Schaller, M.E., Fewell, G.A., Delahunty, K.D., Miner, T.L., Nash, W.E., Cordes, M., Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., Ali, J., Andrews, S., Isak, A., Vanbrunt, A., Nguyen, C., Du, F., Lamar, B., Courtney, L., Kalicki, J., Ozersky, P., Bielicki, L., Scott, K., Holmes, A., Watkins, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C., Dauphin-Kohlberg, S., Kozlowicz-Reilly, A., Leonard, S., Kohlfing, T., Rock, S.M., Tin-Wollam, A.M., Abbott, A., Minx, P., Maupin, R., Strommatt, C., Latreille, P., Miller, N., Johnson, D., Murray, J., Moesner, J.P., Wendt, M.C., Yang, S.P., Schultz, B.R., Wohlmann, P.E., Spieth, J., Bieri, T.A., Nelson, J.O., Berkowicz, N., Wohlmann, P.E., Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Bedell, J.A., Mardis, E.R., Clifton, S.W., Chissoe, S.L., Watta, W.A., Raymond, C., Hugen, E., Gillett, W., Zhou, Y., James, R., Phelps, K., Iadonato, S., Bubb, K., Simms, E., Levy, R., Clendenning, J., Kaul, R., Kent, W.D., Furey, T.S., Baertsch, R.A., Brent, M.R., Keibler, E., Flieck, P., Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrente, D., Chinwalla, A.T., Gish, W.R., Bddy, S.R., McPherson, J.D., Olson, M.V., Eichler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.
The DNA sequence of human chromosome 7
Nature 424 (6945), 157-164 (2003)
22737999
12853948
2 (bases 1 to 154959)
Jones, K., Keppeler, D. and Tin-Wollam, A.
The sequence of Homo sapiens PAC clone RP5-907C10
Unpublished (2001)
3 (bases 1 to 154959)
Waterston, R.H.
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

REFERENCE	MO 63108, USA
AUTHORS	4 (bases 1 to 154959)
TITLE	Waterston, R.H.
JOURNAL	Direct Submission
	Submitted (14-JAN-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 154959)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (22-JAN-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	6 (bases 1 to 154959)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	7 (bases 1 to 154959)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	8 (bases 1 to 154959)
AUTHORS	Wilson, R.
TITLE	Direct Submission
JOURNAL	Submitted (30-JAN-2004) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Jan 14, 1999 this sequence version replaced gi:3213072.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by 'high quality' data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NIGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GTB/CHR7>, send <mailto:green@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION: This clone was derived from human PAC library RPI-5, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., *Nature Genetics* 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong. VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of RP5-907C10
actual end is at 154959 of RP5-907C10.

FEATURES	location/Qualifiers
source	1..154959 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="7" /map="7"
repeat_region	/clone="RP5-907C10" /clone_id="RP5-5" 68..391 /rpt_family="L1" 731..752 /rpt_family="AT_rich"
repeat_region	1128..1429 /rpt_family="Alu" 1554..1648 /rpt_family="L1" 1659..1786 /rpt_family="L1" 1790..1936 /rpt_family="L1"
repeat_region	1949..2296 /rpt_family="L1" 2297..2595 /rpt_family="Alu" 2596..2760 /rpt_family="L1"
repeat_region	2814..26553 /gene="POT1" join(2814..2870,6849..7005,8634..8839,14398..14533,20470..20558,22507..22598,24455..24560,25738..26553)
repeat_region	/gene="POT1" join(2814..2870,6849..7005,8634..8839,14398..14533,20470..20558,22507..22598,24455..24560,25738..26553)
repeat_region	/gene="POT1" join(<2814..2870,6849..7005,8634..8839,14398..14533,20470..20558,22507..22598,24455..24560,25738..26553)
repeat_region	/note="Homo sapiens, clone MGC:10280 IMAGE:3955573, mRNA, complete cds.; H_DJ0907C10.4 This gene was based on gi12804138 13123773) Continues from H_NH056314.1" /codon_start=3 /product="unknown" /protein_id="AAD0852.1" /db_xref="gi:4176376"
repeat_region	/translation="SGSVSLYEVRCCQLSATILTDHQVLETPCLCALIKQAPQYR IRATRSYKPRRLFSQVKHCPKCHLAEVPHEDLDIIFODGATKTPDVAKLQNTSLY DSKTIWTTNQGKRCVAVHYVNNQILPLNSBLLIGSLTSEICKSNKPNYSIVPR SGHDELLELDASAPLLIGCTIHHYGCQKQCSLRITQNLNSLVDTKTSIPSSVNAALGI VPLQVFMPTPLDDGGVLEAYLMDSDKFFQIPASSVLMDDDLQKSVDMIMDMFCPP GIKIDAYPMLECFKSYNVTNMGTDNQICYOIFDTVAEDVI"
repeat_region	3436..4963 /rpt_family="L1" 5057..5355 /rpt_family="Alu" 5384..5685 /rpt_family="Alu" 6265..6308 /rpt_family="CT-rich" 6628..6748 /rpt_family="L1" 7101..7121 /rpt_family="AT_rich" 7421..7446 /rpt_family="AT_rich" 8020..8148 /rpt_family="MIR" 8174..8194 /rpt_family="AT_rich" 9616..9810 /rpt_family="L1" 9834..9908

repeat_region /rpt_family="L1"
9941..10129 /rpt_family="L1"
repeat_region 10420..10528 /rpt_family="L1"
repeat_region 10528..10907 /rpt_family="L1"
repeat_region 10908..11201 /rpt_family="L1"
repeat_region 11202..12388 /rpt_family="L1u"

Query Match 94.4%; Score 17; DB 9; Length 154959;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTATATAG 18
DB 30174 CCAAGCTTTATATAG 30190

RESULT 7
AL133551
LOCUS
DEFINITION
AL133551 175940 bp DNA linear PRI 21-AUG-2000
Human DNA sequence from clone RP11-57G10 on chromosome 10 contains
a J-domain containing protein (JDBP) isoform B, the SIR1 gene
(Sir2-like proteins (sirutinins) type 1), part of a novel gene
similar to KIAA0032, two ribosomal pseudogenes, 2 Cpg islands,
ESTs, STSs and GSSs, complete sequence.

ACCESSION
AL133551.13 GI:9407715
VERSION
HTG; J-domain; RPL21; SIR1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175940)
DIRECT SUBMISSION
Submitted (21-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requester: clonerequests@sanger.ac.uk
On Jul 23, 2000 this sequence version replaced gi:9367364.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-57G10 is from the library RPCI-11.1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-57G10 The true
left end of clone RP11-474D14 is at 172652 in this sequence.

FEATURES
source
1..175940
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"

repeat_region /db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-57G10"
/clone_id="RPCI-11.1"
1..2318
/note="L1P10 repeat: matches 3815..6158 of consensus"
2338..2767
/note="L1M4 repeat: matches 4946..5385 of consensus"
2768..3058
/note="AluLo repeat: matches 1..296 of consensus"
3059..3371
/note="L1M4 repeat: matches 5385..5704 of consensus"
3372..3682
/note="AluSx repeat: matches 1..310 of consensus"
3683..4322
/note="L1M4 repeat: matches 5704..6298 of consensus"
4346..4707
/note="L1M5 repeat: matches 5798..6175 of consensus"
4872..4926
/note="MIR repeat: matches 204..258 of consensus"
4880..4938
/note="L2 repeat: matches 2648..2731 of consensus"
5029..5146
/note="MIR1 repeat: matches 436..564 of consensus"
5150..5194
/note="MIR2CB repeat: matches 2..50 of consensus"
5195..5754
/note="MIR2A repeat: matches 1..453 of consensus"
5755..6125
/note="MIR2CB repeat: matches 42..454 of consensus"
6142..6207
/note="33 copies 2 mer ac 86% conserved"
7403..7494
/note="MIR1 repeat: matches 1..81 of consensus"
7561..7932
/note="MIR1 repeat: matches 180..568 of consensus"
8283..8422
/note="MIR7A repeat: matches 191..336 of consensus"
8415..8907
/note="MIR1A repeat: matches 11..527 of consensus"
8908..9107
/note="MIR7A repeat: matches 1..211 of consensus"
9110..9171
/note="MIR repeat: matches 76..146 of consensus"
9449..9750
/note="AluY repeat: matches 1..304 of consensus"
10489..10938
/note="UTR2 repeat: matches 1..448 of consensus"
11938..12442
/note="Cpg Island"
/evidence="not experimental"
12362..12436
/note="CTRNA-Ser-TCA repeat: matches 8..82 of consensus"
12947..13250
/note="AluY repeat: matches 1..303 of consensus"
13888..13947
/note="MIR1 repeat: matches 453..514 of consensus"
14085..14387
/note="AluSd repeat: matches 1..304 of consensus"
complement(14553..14948)
/note="match: GSS: Em:Q0113044"
14626..14764
/note="MIR1 repeat: matches 229..366 of consensus"
15268..15328
/note="MIR1 repeat: matches 113..172 of consensus"
15822..16120
/note="AluSx repeat: matches 1..306 of consensus"
16150..16463
/note="AluY repeat: matches 1..312 of consensus"
17598..18107
/note="L2 repeat: matches 1843..2330 of consensus"
18101..18225
/note="match: GSS: Em:AQ627100"

misc_feature	18230..18711 /note="match: GSS: Em:AQ627100"
repeat_region	/note="AluX repeat: matches 1..294 of consensus" 19538..19830
repeat_region	/note="AluX repeat: matches 1..294 of consensus" 20049..20256
repeat_region	/note="MER58A repeat: matches 12..224 of consensus" 20724..20880
repeat_region	/note="MIR repeat: matches 20..175 of consensus" 21560..21844
repeat_region	/note="AluX repeat: matches 1..302 of consensus" 21917..22003
misc_feature	/note="L2 repeat: matches 2207..2295 of consensus" 21929..22382
repeat_region	/note="match: GSS: Em:AQ672646" 22135..22178
repeat_region	/note="22 copies 2 mer ac 81% conserved" 22288..22595
repeat_region	/note="AluX repeat: matches 1..305 of consensus" 24457..24579
repeat_region	/note="LTR3 repeat: matches 396..521 of consensus" 24729..24815
repeat_region	/note="LTR3 repeat: matches 125..213 of consensus" 24922..25152
repeat_region	/note="FLAM_A repeat: matches 10..120 of consensus" 25153..25419
repeat_region	/note="AluJo repeat: matches 1..263 of consensus" 25420..25432
repeat_region	/note="FLAM_A repeat: matches 120..133 of consensus" 26538..26802
repeat_region	/note="MER5A repeat: matches 7..189 of consensus" 27004..27072
repeat_region	/note="Alu repeat: matches 238..306 of consensus" 27063..27175
repeat_region	/note="AluJo/PRAM repeat: matches 144..250 of consensus" 27176..27447
repeat_region	/note="AluDb repeat: matches 3..274 of consensus" 27453..27532
repeat_region	/note="L2 repeat: matches 2586..2705 of consensus" 28187..28431
repeat_region	/note="AluDb repeat: matches 30..307 of consensus" 29113..29319
repeat_region	/note="L2 repeat: matches 2476..2693 of consensus" 29643..29942
repeat_region	/note="AluJo repeat: matches 1..293 of consensus" 30116..30180
repeat_region	/note="MLT1C repeat: matches 406..466 of consensus" 30181..30485
repeat_region	/note="AluDb repeat: matches 1..305 of consensus" 30486..30861
repeat_region	/note="MLT1C repeat: matches 1..406 of consensus" 30889..31229
repeat_region	/note="L1WCS repeat: matches 7547..7904 of consensus" 31681..32033
repeat_region	/note="L2 repeat: matches 2379..2709 of consensus" 32163..32593
repeat_region	/note="MER57A repeat: matches 1..434 of consensus" 32634..32976
repeat_region	/note="AluJo repeat: matches 1..298 of consensus" 33377..33567
repeat_region	/note="MER3 repeat: matches 3..201 of consensus" 33583..33734
repeat_region	/note="MER5A repeat: matches 9..189 of consensus" 33759..33812
repeat_region	/note="MER5A repeat: matches 128..181 of consensus" 33813..34125
repeat_region	/note="AluX repeat: matches 1..296 of consensus" 34126..34234
repeat_region	/note="MER5A repeat: matches 21..128 of consensus" 34300..34588
repeat_region	/note="MLT1A repeat: matches 93..361 of consensus" 34899..35075
repeat_region	/note="LTR4 repeat: matches 349..525 of consensus" 35076..35177

Query Match	94.4%	Score 17	DB 9	Length 175940
Best Local Similarity	100.0%	Pred. No. 19		
Matches 17	Conservative 0	Mismatches 0	Indels 0	Gaps 0
<p>Query 1 ACCAAGCTTTATATATA 17 </p>				
<p>Db 81989 ACCAAGCTTTATATATA 82005</p>				
<p>RESULT 8 AY160838/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL</p>				
<p>AY160838 719 bp DNA linear BCT 01-OCT-2003 Uncultured bacterium clone COB Pl-13 16S ribosomal RNA gene, partial sequence. AY160838 AY160838.1 GI:37222277</p>				
<p>uncultured bacterium Bacteria; environmental samples. 1 (bases 1 to 719) Schmitt-Wagner,D., Friedrich,M.W., Wagner,B. and Brune,A. Phylogenetic Diversity, Abundance, and Axial Distribution of Cultures in the Intestinal Tract of Two Soil-Feeding Termites (Cubitermes spp.) Appl. Environ. Microbiol. 69 (10), 6007-6017 (2003)</p>				
<p>2 (bases 1 to 719) Schmitt-Wagner,D., Friedrich,M.W., Wagner,B. and Brune,A. Direct Submission Submitted (09-OCT-2002) Fachbereich Biologie, Mikrobielle Oekologie, Universitaet Konstanz, Universitaetstr. 10, Konstanz D-78457, Germany Location/Qualifiers</p>				
<p>FEATURES source 1..719 /organism="uncultured bacterium" /mol type="genomic DNA" /isolation source="termite intestinal tract" /specific_host="Cubitermes orthognathus, soil-feeding termite" /db_xref="taxon:77133" /clone="COB Pl-13" /environmental_sample <1..>719 /product="16S ribosomal RNA"</p>				
<p>ORIGIN Query Match 88.9% Score 16; DB 1; Length 719; Best Local Similarity 100.0%; Pred. No. 1.3e+02; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0</p>				
<p>Oy 3 CAAGCTTTATATATAG 18 </p>				
<p>Db 49 CAAGCTTTATATATAG 34</p>				
<p>RESULT 9 AB018110/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS</p>				
<p>AB018110 42563 bp DNA linear PLN 14-FEB-2000 Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: K19M13. AB018110 BA000015 AB018110.1 GI:3702728</p>				
<p>Arabidopsis thaliana (chale cress) Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.</p>				
<p>REFERENCE 1 Kaneo,T., Katoh,T., Sato,S., Nakamura,Y., Asanizu,E., Kotani,H.,</p>				

TITLE Miyajima,N. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 5. IX.
Sequence features of the regions of 1,011,550 bp covered by
sevenfold P1 and TAC clones

JOURNAL DNA Res. 6 (3), 183-195 (1999)

MEDLINE 99397451

PubMed 10470850

REFERENCE 2 (bases 1 to 42563)

AUTHORS Nakamura,Y.

TITLE Direct Submission

JOURNAL Submitted (06-OCT-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yata,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)

COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agg_graph.cgi?c=K19M13
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Graal
(Informatics Group, Oak Ridge National Laboratory,
http://combio.ornl.gov/Graal-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-061.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SolicPredictor (Volker Brendel, Stanford University,
http://genemini.zozi.iasitae.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is T32624 and the 3' clone is MQM1.

FEATURES
Source
1..42563
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="K19M13"
/clone_1lb="Mitsui TAC"
/ecotype="Columbia"
2217..3986
/note="gene_id:K19M13.1"
/codon_start=1
/evidence=not_experimental
/product="disease resistance protein-like"
/protein_id="BAB09556.1"
/db_xref="GI:9759078"
/translation="MONLKVYIAMLNLFVSALVRNVLSSGOVICSODRATLLGFKS
SLEDTGVLDVWGKDCNGMGVOCNPATNGKVTGLQSAVNEPILNKKGTLSPS
LGNIRSLIELITENKRTFTGSPNSFSNLTSLRLQIDNLSIQNVSLGHLPLEI
LSLGNRPSGLVPASFGSLRLTNNLARNPSFGIPVTFKNLKLENLDDSSNLISG
PIPIPIQGFQMLNLYLSNRSFGLPVSVSLRLQITMSLERNGLTGPDLSIDRSYK
SLTSLQSGNKFITNPASITGLQIMLSINSLRNFSDPLPVNGARGPSSLIDSYK
NNNLGAIPSWIRDKOLDINLACGKXGTRPKLTPPTTSLDLSNPLFGVAPL
TSLTNNQKXVIRSKNOLRPDLSKTLPGCVASLTLSNLTGSLSLNNKXSPLEET
HLTNNQISGRIPDRGESLNLKVLNIGSKIKSGQIPSSLSNVELYRDLISRNHTTGI
PQATGQAQLKWLSDLSINALTGRIPLDSLINTKTIKXASFRANRLCGQIPQGRPNIFP
AAAYLHNLCICGKPLPACRKTMK"
complement(join(7408..7601,7792..7852))
/note="unnamed protein product; gene_id:K19M13.3
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09557.1"
/db_xref="GI:9759079"
/translation="WVLGLDWTETKHEMESVEDPSKXKPSWRIILNVGPKPKLAWG
HSTCVGSTRVLVIGHNGEEMILNEIHELCLASQSDSL"
complement(join(10326..10490,10573..10656,10748..10792,
10888..10942,11028..11147,11224..11313,11604..11725))
/note="unnamed protein product; gene_id:K19M13.4

unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09558.1"
/db_xref="GI:9759080"
/translation="MAGPSTSNAPKORKEVFEASNTSTTLTRAKDGSAPALCLV
FGCCCCSQAQVETQAEAKKPAEKKTTSSTGPPKRLKTTNDEKSSSTSNKRR
PLTAFPTMSDFRKTPEKSEHNSGLAKDAATIGEEKKSLTEERKYLIDKAELEKYE
NKSLESNDADDEEEDKQSDVDADAEKQVDDDEVEERKEVNTDDDKKEAGKEE
EEELIDY"
complement(join(15556..15709,15838..15906,16013..16129,
16302..16414,16499..16573,16898..17032,17303..17559,
17800..18312,18408..18990,19164..19238,19316..19386,
19825..19927,20466..20692,20785..20921,20999..21140,
21449..21476))
/note="unnamed protein product; contains similarity to
unknown protein
9b|BAB1474.1
gene_id:K19M13.6"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09559.1"
/db_xref="GI:9759081"
/translation="WTKRXYKIOEYFAHSAVNCIKIGKRSRYVLTGGEHKNLM
AIGKNALIVSRVLCQSLYGHSSGIDSVTADASEVLAAGAAGTILKIDLEAK
IVRLTTHRNCISVDPHPPGEPFASGLDNTNLIKIDIRKGCITHKGTGVNVL
FTPDGRNVVGGEDNIIVKVDLTAGKLTLEFKHEGOIQSLDPHPEPLATGSADRT
VKFMDLTPELIGSGEGTAVRCLSPNPGKTVLGLDSEKIFEMEPRIQGDVY
GSRSLSPMNVHEGKLCCSYNOCVGVWVVDLSVPLSYISFELIVAGSNLYEPTH
FQHGQTEECMAGDTQOSNGHBKSCSRDVLVNDNNSKTVLKLSSVQNDPL
KETSILRLSVQNSDSTYKTSIGSSTVSQSSSESKSRFLGRLSVQNSDVKE
SRTPSSGSLPGTPHVSSTNVKATGSGTAVSNATSRNFTKANPKNVAKAD
FAPVIVRADRIEQAATESRAELDIARTPYVLQADSRSSSNNDPLPASYLE
MSEQIPVEPNPIPFGTLPFGKGMKRTATERSINDRKYKSGNSRMSGSPNHD
ENVYLVHRENRDPSPTESQGGDSASQSLFLOKMRISFMCLISGFPOSIVIRE
RGRFNSFEPSPVNSFGNMPARIRSNMFKRGKHMPEQGIIDSPSENYEDIME
KKNQVSSWQSRKLKQVRRVWRNDVKSISIEKMDNAVTAUVDGLITERNITL
TLNDCTSLPLTALAGSGMDHLSVLDLLKLVLYSGPIYSLSAPASGVDDIEA
EORIERYSRCFVELKEVKACLPSLISGSLIITMSLTTLDMNAEQKVFVLEBPAS
SOVPHLLPVDVLPKGV"
complement(24844..25392)
/note="unnamed protein product; contains similarity to
ferredoxin-thioredoxin reductase variable chain
gene_id:K19M13.7"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09560.1"
/db_xref="GI:9759082"
/translation="WSSQIALSPALAAIRPSSHDCLASATTAATPMALKCTIVA
PLSLFTSQSQIKHSSSRKTSRTTIRCDVALTKSDSINADNPSSSPSSSEIEAEKA
KIGSRVVTAPLKYHVNRVPEVDEMEGKLKDYVAVWKGKISANLPIKEFFKEI
EGRLVYFVHLKEDDEFIDQ"
complement(join(26890..27881,28441..28540,28657..28803,
29264..29301,29417..29510,29664..30038,30122..30574))
/note="unnamed protein product; gene_id:K19M13.8
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09561.1"
/db_xref="GI:9759083"
/translation="WQKSGVNRNPSLKVAILPQAQSLRLRLGCSQIATGSGQSSPIV
FPEKRNKVVASSRGEVNDPQKVPDHRIDIGGDEKSLDLSLVAAGLVLDK
RKSAGKDATEIIOOPATDTSIKKAVDAKLTSLAVGSDMLQNLNVSVTVVGLRH
FTVAVYPIGKSGSLGCFTEPKRSDRQFRVATVEBAVQWVASPDQOCFLNCLP
LVAKQASSELFSVPIDTPPELPLFRCKSAKMLVIINPSGHRGSIKVPHNVVEPIEX
LAGIKMEVNTKTAGHRELASTVDINLSCDGIITCGDGIINEVANGLLTNRNPKG
VSLPIGIVPAGSDSLVTVLGVDPISALSLVKGLATVDVFAVEWHTGIIHGM
TVSYGVFVSDVLESEKYQKRFGLRPLVFGKLFKCLPYTSVFEVLLPAQKDAKX
IRLEKEDMDQDLYTVDRSSREGFPASLSISIMPSVGEGLDSCSTASTEP
SEVYRGIDPMKRLSGSRBDVTAPEXVHQAOQSTIPNMRTEKSKRMKGHMLVSV
ODPPTRSKMGNTGGQDREDSSTVSDPGPIWMDGPKMDPESAMVDENSETLDPGRED
LETGLRQOSTIPIPEDVWRSKHFPLGINVGNACRTVQSSQVAVANSEHDDGTMDML
LVHGCGRLRLRPFILLQOTGRHLSLPYECVKVNYLSYFLTLNRLQLANC"

CDS
34686..35021
/note="unnamed protein product; gene_id:K19M13.9
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB09562.1"
/db_xref="GI:9759084"
/translation="MPEQGFYININGASDASPKALIGETLKQFPYSNPEETRC
GKLRDRHREERFTSPQSRLLTVADRNFCGVNPVNPRLNVMTRKLTSLHLVDTKS
TCVSCSLF"
35475..36383
/note="unnamed protein product; gb|AAC78512.1
gene_id:K19M13.10
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB09563.1"
/db_xref="GI:9759085"
/translation="MSLIVETLFLFEPQSKRKINPENTIESPNCNMFVNRKICISY
TVYDNFEGRSFDLYKGMQNSNLTGKLTISQTSMLDEKHLVLDQTLIHTIK
TSLVSEKRIIEVSRKDIKRNFTGPRESLIKRPVHQFLKCNEMFTSVYTK
GGYDVALVLEMDPDKFTYGNRVITPRBSPGKTDLVADRGIVIVDTTSVMPH
DKNLIQIARYKFKGSCLEFSDKKIDSDERKPLMTALRFKDVHEEFYDMSK
DLSVVRPLKELSLRWKRGV"
join(37928..38528,38657..38766,39049..39108,39289..39339,
39675..40085,40152..40239,40345..40769)
/note="unnamed protein product; emb|CAB62634.1
gene_id:K19M13.11
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB09564.1"
/db_xref="GI:9759086"
/translation="MEDNKEKVGSSKKRLRKPKSLFVGMGSRNLIIEFLSGRDT
NKTSENDVATLIMVYREKRETPKSKRRKTVACDEKRLLEGAQINVIKPDIV
EKIYVQEDELPTDDIYASDDKQQLSLSDYAKOTKQVSKRGRTFAIRIVTVK
LVLYKSLVGLATPETFESKVRITVRINKPQLVHVTGVKSGDIDGNLQVYNY
SVYIKDVTSLDDDFSOBCEELHQRINNGFAKRLTVDMESKASLHEDVTKHL
EKRELQNPDEQKLVDEPEIYVAELEPECEDDDDRTIEDLIIVNPEAHOSDKQ
RORDLPVSSVYKQKQENLLKNPEBOLRLCDVDPVAEKLPEFVDDGKLVNDVP
YQTKPSLRITKWKYKNQVRLICNNSIGLKLVLGNPDIRPCLQSNKMDLHEDVY
BPPANGITLMTDSTTBGEMNTKVSQHSSTPVIDLSKTKQASHNPIIILSDDDDD

Query Match 88.9%; Score 16; DB 8; Length 42563;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTAAATA 17
|||||
Db 33745 CCAAGCTTTTAAATA 33730

RESULT 10
AF467900/c 48843 bp DNA linear PLN 22-MAY-2003
LOCUS
DEFINITION Prunus persica clone pbn31C7 hypothetical protein, and
hypothetical transcription factor genes, complete cds.
AF467900
ACCESSION
VERSION AF467900.1 GI:27450528
KEYWORDS
SOURCE
ORGANISM Prunus persica (peach)
Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosida I; Rosales; Rosaceae; Amygdaloideae; Prunus.
1 (baes 1 to 48843)
Georgi, L., Wang, Y., Yvergniaux, D., Ormsbee, T., Inigo, M.,
Reignard, G., and Abbott, G.
Construction of a BAC library and its application to the
identification of simple sequence repeats in peach [Prunus persica
(L.) Batsch]
Theor. Appl. Genet. 105 (8), 1151-1158 (2002)
JOURNAL
PUBMED 12582893

REFERENCE 2 (baes 1 to 48843)
AUTHORS Georgi, L., Wang, Y., and Abbott, A. G.
TITLE Direct Substitution
JOURNAL Submitted (11-JAN-2002) Genetics and Biochemistry, Clemson
University, 122 Long Hall, Clemson, SC 29634, USA
FEATURES
source
1..48843
/organism="Prunus persica"
/mol_type="genomic DNA"
/culivar="hemared"
/db_xref="taxon:3760"
/clone="pbn31C7"
/issue_type="rootstock"
/clone_lib="pbelobac11"
/note="derived from genomic DNA partially digested with
Hind III"
291..318
/note="similar to a direct repeat on Arabidopsis thaliana
chromosome 2"
/tpc_type="inverted"
355..393
/note="similar to a direct repeat on Arabidopsis thaliana
chromosome 2"
/tpc_type="inverted"
join(3833..3907,4079..4288,4418..4542,4630..5560)
/product="hypothetical protein"
join(3833..3907,4079..4288,4418..4542,4630..5560)
/note="similar to hypothetical protein F6P9.9 in
Arabidopsis thaliana"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA014624.1"
/db_xref="GI:27450529"
/translation="MGNGRNALICHTVYSRSLIKISVYKFLIERSPOVGLASOD
HLQAKSWLSHSGTSDILPFGFAGVANSQLOQVNSQIPILIKRCPPRYVLTPT
QVCGESEKEVEIONREIRLEAVYPRSAIPLNSVLADVEDSGHNDVQVPVIT
SIDDEGVGDTSSASMAPNHTSTOSFLSVQGLAASQSVVNNRNHPTKSNVAGI
PUGVEVDVYAAAGATVNSNHNMDIELIKLSNPKMIEKLVKQPPQVTRP
OMGFIADPPPHNIRTSSTSRSSSTALSGHFYQPVAVGMRPPDAPPPHAAISVS
PPVAGPPPADIVYKSLIQDGGDHRDPRPPGNRSHSALINOSNNKYSKDSKP
KIMKPCIFNYSRGRHGANCAVQHASFPQSSSAPEVHSTRKMDREISS"
join(46350..6706,7104..7133)
/product="hypothetical protein"
join(6350..6706,7104..7133)
/note="similar to unknown protein MJ24.9 in Arabidopsis
thaliana"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA014625.1"
/db_xref="GI:27450530"
/translation="MLKFLMRVRIERPADPRTASCLEFLAOCNARAKXSNPSCAVO
VKRRTDEPKRTVTVFVNGVEVADATATPAQDRIKMTLEKQSLTEQVFRDGEPM
PVIIIPAEELHQPAGTKPRKADERKO"
complement(join(7784..7882,7973..8098,8313..8512,
8627..8865,9253..9440,9712..10323,10419..10574))
/product="hypothetical protein"
complement(join(7784..7882,7973..8098,8313..8512,
8627..8865,9253..9440,9712..10323,10419..10574))
/note="similar to unknown protein MJ24.8 in Arabidopsis
thaliana"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA014626.1"
/db_xref="GI:27450531"

misc_feature
join(12687, .15371,15371, .15549,16324, .16402,16405, .16717)
/note="similar to putative DEAH helicase MPE16.9 in Arabidopsis thaliana; likely non-functional due to two frame-shifts and premature stop codons"
<17499, .18764
/product="hypothetical protein"
17499, .18764
/note="similar to unknown protein K17E12.5 in Arabidopsis thaliana"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA014627.1"
/db_xref="GI:27450532"
/translation="MGSVSLKIGDGTAFKRAATLCSAVMLMLFVITTNLFALYAF
TSPPDOQYHLHHTQKNIISLSEVSLIRLSEDSQKLAQKELGYESIDSR
SNVAHELKLFLOHLOPLGKDSRGITGVMSVAVSOSKDLISOMYKSGPCD
DMSLAQKILRGCERPRRCRAATLPKVGINPPRILMKPVSGKITVSGLSKSE
CLNSKLSRDCVGFDLVNGENORFYKAKNDPLIDVLAISGGITIGFDIGGS
GTPAARMAERNMTVITNTLNDAPSEFIFARGLPFLSDPPYDVNFDLVHAA
SGLDVGGKPKPEFEMEDIRLIRPGGLFIMDNYCSNEEKRLDITLIERFGYKXK
WVVDKVDAAAGSEVYLAVLOKPVAV"
complement(join(19604, .19745,19833, .20023,20115, .21436,
21513, .21648,21782, .21857,22120, .22242,22363, .22527,
22823, .22913,23088, .23172,23278, .23430,23534, .23590,
23701, .23799,24052, .24164,24372, .24483))
/product="hypothetical transcription factor"
complement(join(19604, .19745,19833, .20023,20115, .21436,
21513, .21648,21782, .21857,22120, .22242,22363, .22527,
22823, .22913,23088, .23172,23278, .23430,23534, .23590,
23701, .23799,24052, .24164,24372, .24483))
/note="similar to IAA24/monoperox/Arf5 gene, F6f9.10
protein, in Arabidopsis thaliana"
/codon_start=1
/product="hypothetical transcription factor"
/protein_id="AA014628.1"
/db_xref="GI:27450533"
/translation="MINGSVEEIKKAGGLSGAQSILDEKMLKEIQDSGSKRAIN
SELMACAGLVLCIPVQSLSYPPQHSQVAVSTRKATSGIPNPNLPSQLCV
ONVTLHADKETDEIYAQMSLKPNSKVDVPPVFGKPKSGHPFCKTLTADST
HGGSFVPRARAEKLPRLDTPMOPSGEIVRDLHDSMFRHVRPOPKHLLTGM
SLFVGAKRLAAGSVLFRDEKSLMGRRANQQTTPBSVLSAASHIGVUAAA
HAANRSPTFTFNPACPSFVPLATYQALITGDSVSMRGMPETESKRRY
MGITVSTSDLPDLPWPSKWRNLQVDEPCCQKONRVSMELETENPIFPLSLTS
SLRPSHTGFLGATEWGNLIRKPIFVPEIGNNSFPYSISNLSQSLVMLKPOL
VNHAGTALAQOOSFANGDLADKMAQAKLQKNPVSFSGTSLQSNPQSLDSS
ATIDVNTTSHAILPKLNLTKESQAPVGNSTDKLETDFPSADLSQNSLGLGE
DKLAAGVSRYNLNLTFRANQSAALQTSPPMPQPLSLIYHQQDPMPSDN
STNGSLFLNDCTFYQSTQPFAGTIRSQGSLVFGQDSAVLLEANNSSLTISIG
EMWNSLNLNRLLPQVDLTLSHQGSLNCSLSSSLRSLDSNNQSGIYGCYNDV
VSGSVAVIDPVSSTILDFSTLKNADFNHPSDCLGNLSSQDLQSLTASLIGDS
QAFSRDLADNSGSSSNIDLDESSLONNNSHWQVPPRYTKYQKTSVGRSID
VPSFKYVEELCSAIECMFGLGLNDPRGSGMKLVYVDYDVLVDDDPHEEVGCV
RCRILSPTEVQMSSEGITKLNANMGINGTMSSEGRY"
join(36926, .36970,36973, .37377)
/note="similar to hypothetical protein F6f9.11 in Arabidopsis thaliana; non-functional due to a frame-shift"
41043, .41144
repeat_region
41144, .41437
/note="likely the remains of a retroelement"
repeat_region
41144, .41437
/note="likely the remains of a retroelement"
41789, .42253
/note="likely the remains of a retroelement"
repeat_region
42253, .42410
/note="likely the remains of a retroelement"

Db 5275 CAAGCTTTATATAG 5260
RESULT 11
AC145628
LOCUS
DEFINITION
Homo sapiens chromosome UNK clone XFOS-802760B8, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
ACCESSION
AC145628
VERSION
AC145628.1 GI:32996967
KEYWORDS
HTG: HTGS PHASE1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 51797)
Wilson, R.K.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 51797)
Wilson, R.K.
Direct Submission
Submitted (19-JUL-2003) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project information

Center project name: H Aa802760B8

Sequencing vector: pBluescript
Sequencing method: Sanger
Chemistry: Dye-terminator Big Dye, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 42380 bases at least Q40
Consensus quality: 44639 bases at least Q30
Consensus quality: 46559 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 3260: contig of 3260 bp in length
*
3261 3360: gap of unknown length
*
3361 7533: contig of 4172 bp in length
*
7533 7633: gap of unknown length
*
7633 12298: contig of 4666 bp in length
*
12298 12399: gap of unknown length
*
12399 21234: contig of 8836 bp in length
*
21234 21334: gap of unknown length
*
21334 46167: contig of 24833 bp in length
*
46167 46268: gap of unknown length
*
46268 47548: contig of 1281 bp in length
*
47548 47649: gap of unknown length
*
47649 49886: contig of 2238 bp in length
*
49886 49987: gap of unknown length
*
49987 51797: contig of 1811 bp in length.
*
Location/Qualifiers
1. 51797
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UNK"
/clone="XFOS-802760B8"
1. 3260

Query Match 88.9%; Score 16; DB 8; Length 48843;
Best local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
3 CAAGCTTTATATAG 18
|||||

misc_feature
1. 3260


```

misc_feature /note="assembly_name:Contig10"
3361..7532
/note="assembly_name:Contig11"
7633..12298
misc_feature /note="assembly_name:Contig12"
12399..21234
misc_feature /note="assembly_name:Contig13"
21335..46167
/note="assembly_name:Contig14"
46268..47548
misc_feature /note="assembly_name:Contig6"
47649..49886
/note="assembly_name:Contig8"
49987..51797
misc_feature /note="assembly_name:Contig9"

```

ORIGIN

```

Query Match      88.9%; Score 16; DB 2; Length 51797;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3  CAACTTTTATATAG 18
Db      42039 CAACTTTTATATAG 42054

```

RESULT 12

```

AC109350 58068 bp DNA linear PRI 30-MAR-2002
LOCUS AC109350
DEFINITION Homo sapiens BAC clone RP11-267D2 from 4, complete sequence.
ACCESSION AC109350
VERSION AC109350.5 GI:19526559
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 58068)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PMID 9847074

```

```

REFERENCE 2 (bases 1 to 58068)
AUTHORS Buatesi,D. and Cotton,M.
TITLE The sequence of Homo sapiens BAC clone RP11-267D2
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 58068)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

REFERENCE 4 (bases 1 to 58068)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

REFERENCE 5 (bases 1 to 58068)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

REFERENCE 6 (bases 1 to 58068)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

```

```

COMMENT On Mar 17, 2002 this sequence version replaced gi:19339131.
-----Genome Center
Center: Washington University Genome Sequencing Center

```

```

Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
-----Summary Statistics
Center project name: H_NH0267D02
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenio,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-609N18, 2000 bp overlap; the clone sequenced to the right is CTD-2012117, 2000 bp overlap. Actual end of this clone is at base position 105788 of CTD-2012117.

Polymorphisms exist between AC093722 and AC109350.

FEATURES

source

```

1..58068
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-267D2"
/clone_lib="RPCT-11"
423..462
/rpc_family="AT_rich"
1124..1158
/rpc_family="AT_rich"
1283..1310
/rpc_family="AT_rich"
1753..1809
/rpc_family="AT_rich"
2483..2513
/rpc_family="AT_rich"
3343..3373
/rpc_family="AT_rich"
3694..3714
/rpc_family="(GA)n"
5020..5146
/rpc_family="AT_rich"
5203..5251
/rpc_family="MIR"
5524..5626
/rpc_family="(CAT)n"
5666..5954
/rpc_family="L2"
/rpc_family="Alu"

```

```

repeat_region      6388..6429
                    /rpt_family="(CA)n"
repeat_region      6741..7042
                    /rpt_family="Alu"
repeat_region      7138..7697
                    /rpt_family="ERV1"
repeat_region      8386..8452
                    /rpt_family="(CATTA)n"
repeat_region      8635..9011
                    /rpt_family="MALR"
repeat_region      9418..9439
                    /rpt_family="AT_rich"
repeat_region      9633..9654
                    /rpt_family="AT_rich"
repeat_region      9696..10177
                    /rpt_family="AT_rich"
repeat_region      10178..10461
                    /rpt_family="Alu"
repeat_region      10483..10516
                    /rpt_family="AT_rich"
repeat_region      10599..11288
                    /rpt_family="ERV1"
repeat_region      11707..11741
                    /rpt_family="AT_rich"
repeat_region      12781..13130
                    /rpt_family="ERV1"
repeat_region      13131..13276
                    /rpt_family="Alu"
repeat_region      13277..13561
                    /rpt_family="ERV1"
repeat_region      13776..13796
                    /rpt_family="AT_rich"
repeat_region      13937..13957
                    /rpt_family="AT_rich"
repeat_region      14622..14853
                    /rpt_family="AT_rich"
repeat_region      14854..15181
                    /rpt_family="Alu"
repeat_region      15182..15212
                    /rpt_family="ERV1"
repeat_region      15199..15424
                    /rpt_family="ERV1"
repeat_region      15425..15766
                    /rpt_family="ERV1"
repeat_region      15767..15912
                    /rpt_family="ERV1"
repeat_region      16747..16783
                    /rpt_family="L2"
repeat_region      17475..17545
                    /rpt_family="AT_rich"
repeat_region      17721..17806
                    /rpt_family="L2"
repeat_region      17874..17982
                    /rpt_family="MER1_type"
repeat_region      18058..18078
                    /rpt_family="AT_rich"
repeat_region      18196..18472
                    /rpt_family="Alu"
repeat_region      18620..18705
                    /rpt_family="MIR"
repeat_region      18706..19022
                    /rpt_family="Alu"
repeat_region      19023..19110
                    /rpt_family="MIR"
repeat_region      19159..19318
                    /rpt_family="L2"
repeat_region      19626..19685
                    /rpt_family="CT-rich"
repeat_region      19726..19787
                    /rpt_family="Alu"
repeat_region      19964..20099
                    /rpt_family="ERV1"
repeat_region      20419..20763

```

```

repeat_region      20765..21143
                    /rpt_family="L1"
repeat_region      21349..21390
                    /rpt_family="AT_rich"
repeat_region      21478..22305
                    /rpt_family="L1"
repeat_region      22306..22597
                    /rpt_family="Alu"

Query Match      88.9%; Score 16; DB 9; Length 58068;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      2 CCAAGCTTTATATATA 17
Db      56067 CCAAGCTTTATATATA 56082

RESULT 13
AC092284/c      107044 bp      DNA      linear      PRI 26-FEB-2002
LOCUS      Homo sapiens chromosome 5 clone CTD-2160D9, complete sequence.
ACCESSION      AC092284
VERSION      AC092284.2 GI:18921294
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS      1 (bases 1 to 107044)
TITLE      DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL      Direct Submission
REFERENCE
AUTHORS      2 (bases 1 to 107044)
TITLE      DOE Joint Genome Institute.
JOURNAL      Direct Submission
AUTHORS      Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint
JOURNAL      Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS      3 (bases 1 to 107044)
TITLE      DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL      Direct Submission
AUTHORS      Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
JOURNAL      Drive, Walnut Creek, CA 94598, USA
COMMENT      On Feb 26, 2002 this sequence version replaced gi:14589473.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

FEATURES
source
1..107044
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2160D9"

ORIGIN

Query Match      88.9%; Score 16; DB 9; Length 107044;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      3 CAAGCTTTATATATG 18
Db      68352 CAAGCTTTATATATG 68337

RESULT 14
AC092359      107129 bp      DNA      linear      HTG 07-MAR-2002
LOCUS      Homo sapiens chromosome 5 clone RP11-341E24, WORKING DRAFT
DEFINITION

```

ACCESSION	SEQUENCE, 3 ordered pieces.
VERSION	AC092359
KEYWORDS	AC092359.2 GI:1922482
SOURCE	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 107129)
TITLE	DOE Joint Genome Institute.
JOURNAL	Sequencing of Human Chromosome 5
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 107129)
TITLE	DOE Joint Genome Institute.
JOURNAL	Direct Submisison
REFERENCE	Submitted (03-JUN-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS	3 (bases 1 to 107129)
TITLE	DOE Joint Genome Institute.
JOURNAL	Direct Submisison
REFERENCE	Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Mar 7, 2002 this sequence version replaced gi:14589548.

```
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 537429
Center clone name: RPCI-11_341B24
```

```

Summary Statistics
Consensus quality: 104726 bases at least Q40
Consensus quality: 105971 bases at least Q30
Consensus quality: 106638 bases at least Q20
Estimated insert size: 161520; agarose-fp estimation
Estimated insert size: 166939; sum-of-contigs estimation
Quality coverage: 6.12 in Q20 bases; agarose-fp estimation
Quality coverage: 10.39 in Q20 bases; sum-of-contigs estimation
NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  1 27924: contig of 27924 bp in length
    27925 28024: gap of unknown length
    28025 29265: contig of 1245 bp in length
    29270 29365: gap of unknown length
    29370 107129: contig of 77760 bp in length.

```

FEATURES

Source

ORIGIN

Query Match	88.9%	Score 16;	DB 2;	Length 107129;
Best Local Similarity	100.0%;	Pred. No. 74;		
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	3	CAAGCTTTTATAATAG	18
Db	68444	CAAGCTTTTATAATAG	68429

RESULT 15	
LOCUS	AC104793/c
DEFINITION	AC104793 bp DNA linear PRI 16-APR-2002
ACCESSION	Homo sapiens BAC clone RP11-23406 from 4, complete sequence.
VERSION	AC104793.4 GI:19909440
KEYWORDS	HTG.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	Sulston,J.E. and Waterston,R.
JOURNAL	Toward a complete human genome sequence
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)
PUBMED	99063792
REFERENCE	9847074
AUTHORS	2 (bases 1 to 133244)
TITLE	Swearengen-Shahid,S., Kozlowski,A. and Dignan,G.
JOURNAL	The sequence of Homo sapiens BAC clone RP11-23406
REFERENCE	Unpublished (2001)
AUTHORS	3 (bases 1 to 133244)
TITLE	Waterston,R.H.
JOURNAL	Direct Submission
REFERENCE	Submitted (21-DEC-2001) Genome Sequencing Center, Washington
AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
TITLE	MO 63108, USA
JOURNAL	4 (bases 1 to 133244)
REFERENCE	Waterston,R.H.
AUTHORS	Direct Submission
TITLE	Submitted (11-MAR-2002) Genome Sequencing Center, Washington
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
REFERENCE	MO 63108, USA
AUTHORS	5 (bases 1 to 133244)
TITLE	Waterston,R.H.
JOURNAL	Direct Submission
REFERENCE	Submitted (03-APR-2002) Genome Sequencing Center, Washington
AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
TITLE	MO 63108, USA
JOURNAL	6 (bases 1 to 133244)
REFERENCE	Waterston,R.
AUTHORS	Direct Submission
TITLE	Submitted (16-APR-2002) Department of Genetics, Washington
JOURNAL	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Apr 3, 2002 this sequence version replaced gi.19339126.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPc1-1 human BAC library was made from the blood of one male donor, as described by Oseegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateo, M., Cataneese, J. J., and de Jong, P. J. (1996) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-80B23, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-23406; actual end is at base position 3275 of RP11-80B23.

The region from 17296 to 17363 is covered only by a PCR product from clone DNA.

FEATURES

NRs	Location/Qualifiers
source	1..133244
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-23406"
	/clone_1fb="RPC1-11"
repeat_region	1..279
	/rpt_family="MaIR"
repeat_region	899..977
	/rpt_family="I2"
repeat_region	2465..2496
	/rpt_family="AT_rich"
repeat_region	2506..2709
	/rpt_family="ERVH"
repeat_region	2710..3136
	/rpt_family="MaIR"
repeat_region	3137..3372
	/rpt_family="ERVH"
repeat_region	3794..3923
	/rpt_family="MIR"
repeat_region	4072..4222
	/rpt_family="ERVH"
repeat_region	4274..4710
	/rpt_family="ERVH"
repeat_region	5294..5377
	/rpt_family="ERVH"
repeat_region	5477..5732
	/rpt_family="ERVH"
repeat_region	5953..6636
	/rpt_family="ERVH"
repeat_region	6726..7020
	/rpt_family="Alu"
repeat_region	8017..8117
	/rpt_family=" (TA) n"
repeat_region	8599..8631
	/rpt_family="AT_rich"
repeat_region	8674..8924
	/rpt_family="ERVH"
repeat_region	8931..8906
	/rpt_family="ERVH"
repeat_region	9791..9891
	/rpt_family="ERVH"
repeat_region	9951..10767
	/rpt_family="ERVH"
repeat_region	10769..11282
	/rpt_family="L1"
repeat_region	11286..11594
	/rpt_family="Alu"
repeat_region	11597..11729
	/rpt_family="ERVH"
repeat_region	11866..11693
	/rpt_family=" (TA) n"
repeat_region	11901..12013

```
repeat_region /rpt_family="ERV1" 12853..12922
repeat_region /rpt_family="ERV1" 12923..13229
repeat_region /rpt_family="Alu" 13230..13606
repeat_region /rpt_family="ERV1" 14053..14471
repeat_region /rpt_family="MALR" 14537..15128
repeat_region /rpt_family="ERV1" 15255..15553
repeat_region /rpt_family="Alu" 15723..15743
repeat_region /rpt_family="AT_rich" 15987..16007
repeat_region /rpt_family="AT_rich" 16386..16548
repeat_region /rpt_family="MIR" 16557..17146
repeat_region /rpt_family="ERV1" 17170..17363
repeat_region /rpt_family="(TTATG)n" 17365..17498
repeat_region /rpt_family="Alu" 18610..18796
repeat_region /rpt_family="MIR" 18801..18823
repeat_region /rpt_family="AT_rich" 20968..21013
repeat_region /rpt_family="AT_rich" 21403..21438
repeat_region /rpt_family="(TTTTG)n" 22435..24301
repeat_region /rpt_family="L1" 24299..24540
repeat_region /rpt_family="L1" 24621..24910
repeat_region /rpt_family="Alu" 25336..25371
repeat_region /rpt_family="(TG)n" 26548..26849
repeat_region /rpt_family="Alu" 27041..27080
repeat_region /rpt_family="AT_rich" 27135..28394
repeat_region /rpt_family="L2" 28401..28761
repeat_region /rpt_family="MALR" 28781..29106
repeat_region /rpt_family="L2" 30336..30708
repeat_region /rpt_family="MALR" 31106..31201
repeat_region /rpt_family="MIR" 31600..31770
```

Query March	88.9%	Score 16	DB 9	Length 13344
Best Local Similarity	100.0%	Pred. No. 72		
Matches 16	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	3	CAAGCTTTTAAATAG	18	
Db	111702	CAAGCTTTTAAATAG	11687	

Search completed: February 10, 2005, 13:14:00
Job time : 117.027 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 5.5316 Seconds
(Without alignments)
5324.730 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18

Sequence: 1 accaagctttatataag 18

Scoring table: OLIGO NUC
Gapex 60.0, Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Patente NA.*
1: /cgm2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgm2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgm2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	4	US-09-872-051-12
2	18	100.0	1183	4	US-09-872-051-8
3	17	94.4	45587	4	US-09-949-016-15836
4	15	83.3	601	4	US-09-949-016-129275
5	15	83.3	601	4	US-09-949-016-129346
6	15	83.3	601	4	US-09-949-016-129417
7	15	83.3	601	4	US-09-949-016-129488
8	15	83.3	601	4	US-09-949-016-129559
9	15	83.3	601	4	US-09-949-016-129630
10	15	83.3	601	4	US-09-949-016-129701
11	15	83.3	601	4	US-09-949-016-129772
12	15	83.3	601	4	US-09-949-016-129843
13	15	83.3	601	4	US-09-949-016-129914
14	15	83.3	601	4	US-09-949-016-129985
15	15	83.3	601	4	US-09-949-016-130056
16	15	83.3	2152	4	US-09-449-632-3
17	15	83.3	107329	4	US-09-949-016-12663
18	15	83.3	107329	4	US-09-949-016-12664
19	15	83.3	107330	4	US-09-949-016-15408
20	15	83.3	107330	4	US-09-949-016-15409
21	15	83.3	107330	4	US-09-949-016-15410
22	15	83.3	107330	4	US-09-949-016-15411
23	15	83.3	107750	4	US-09-949-016-12662
24	15	83.3	107751	4	US-09-949-016-15412
25	15	83.3	107751	4	US-09-949-016-15413
26	15	83.3	107751	4	US-09-949-016-15414
27	15	83.3	107751	4	US-09-949-016-15415

28	15	83.3	107925	4	US-09-949-016-11875	Sequence 11875, A
29	15	83.3	107926	4	US-09-949-016-15404	Sequence 15404, A
30	15	83.3	107926	4	US-09-949-016-15405	Sequence 15405, A
31	15	83.3	107926	4	US-09-949-016-15406	Sequence 15406, A
32	15	83.3	107926	4	US-09-949-016-15407	Sequence 15407, A
33	15	83.3	107926	4	US-09-949-016-15408	Sequence 15408, A
34	15	83.3	107926	4	US-09-949-016-15409	Sequence 15409, A
35	15	83.3	107926	4	US-09-949-016-15410	Sequence 15410, A
36	15	83.3	107926	4	US-09-949-016-15411	Sequence 15411, A
37	15	83.3	107926	4	US-09-949-016-15412	Sequence 15412, A
38	15	83.3	107926	4	US-09-949-016-15413	Sequence 15413, A
39	15	83.3	107926	4	US-09-949-016-15414	Sequence 15414, A
40	15	83.3	107926	4	US-09-949-016-15415	Sequence 15415, A
41	15	83.3	107926	4	US-09-949-016-15416	Sequence 15416, A
42	15	83.3	107926	4	US-09-949-016-15417	Sequence 15417, A
43	15	83.3	107926	4	US-09-949-016-15418	Sequence 15418, A
44	15	83.3	107926	4	US-09-949-016-15419	Sequence 15419, A
45	15	83.3	107926	4	US-09-949-016-15420	Sequence 15420, A

ALIGNMENTS

RESULT 1
US-09-872-051-12
Sequence 12, Application US/09872051
Patent No. 6825400
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMGT32(Ink603) and Composition and Methods for Det
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(18)
OTHER INFORMATION: Agrobacterium tumefaciens nos 3' terminator and rice actin prom
OTHER INFORMATION: et DN
US-09-872-051-12
Query Match 100.0%; Score 18; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
C
1 ACCTGCTTTATATATAG 18
1 ACCTGCTTTATATATAG 18
RESULT 2
US-09-872-051-8
Sequence 8, Application US/09872051
Patent No. 6825400
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl

```
/ APPLICANT: Hironaka, Catherine
/ APPLICANT: Heck, Gregory
/ APPLICANT: You, Jinsong
/ TITLE OF INVENTION: Corn Event PV-ZMG132 (nk603) and Composition and Methods for Detect
/ TITLE OF INVENTION: Thereof
/ FILE REFERENCE: 38-21(52258)B
/ CURRENT APPLICATION NUMBER: US/09/872,051
/ PRIOR FILING DATE: 2001-06-01
/ PRIOR APPLICATION NUMBER: 60/213,567
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 60/241,215
/ PRIOR FILING DATE: 2000-10-13
/ PRIOR APPLICATION NUMBER: 60/240,014
/ PRIOR FILING DATE: 2000-10-13
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 8
/ LENGTH: 1183
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ NAME/KEY: source
/ LOCATION: (1)..(1183)
/ OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
/ OTHER INFORMATION: 165-381 construct vector DNA
/ OTHER INFORMATION: 382-686 Zea mays plastid genes, rps11 and rpoA
/ OTHER INFORMATION: 687-1183 Zea mays genomic DNA
US-09-872-051-8
```

```
Query Match          100.0%; Score 18; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1  ACCAAGCTTTATATAG 18
DB      156  ACCAAGCTTTATATAG 173
```

```
RESULT 3
US-09-949-016-15836/c
/ Sequence 15836, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15836
/ LENGTH: 45587
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-15836
```

```
Query Match          94.4%; Score 17; DB 4; Length 45587;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1  ACCAAGCTTTATATA 17
DB      6030  ACCAAGCTTTATATA 6014
```

RESULT 4

```
US-09-949-016-129275
/ Sequence 129275, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 129275
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-129275
```

```
Query Match          83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2  CCAAGCTTTATAT 16
DB      548  CCAAGCTTTATAT 562
```

```
RESULT 5
US-09-949-016-129346
/ Sequence 129346, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 129346
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-129346
```

```
Query Match          83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2  CCAAGCTTTATAT 16
DB      548  CCAAGCTTTATAT 562
```

```
RESULT 6
US-09-949-016-129417
/ Sequence 129417, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129417
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129417

Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCAAGCTTTTAAAT 16
DB      548 CCAAGCTTTTAAAT 562

RESULT 7
US-09-949-016-129488
; Sequence 129488, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129488
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129488

Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCAAGCTTTTAAAT 16
DB      548 CCAAGCTTTTAAAT 562

RESULT 8
US-09-949-016-129559
; Sequence 129559, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
```

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129559
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129559

Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCAAGCTTTTAAAT 16
DB      548 CCAAGCTTTTAAAT 562

RESULT 9
US-09-949-016-129630
; Sequence 129630, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129630
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129630

Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCAAGCTTTTAAAT 16
DB      548 CCAAGCTTTTAAAT 562

RESULT 10
US-09-949-016-129701
; Sequence 129701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 129701
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-129701

Query Match 83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
|||||
Db 548 CCAAGCTTTTATAT 562

RESULT 11
US-09-949-016-129772
;; Sequence 129772; Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CI001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 129772
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-129772

Query Match 83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
|||||
Db 548 CCAAGCTTTTATAT 562

RESULT 12
US-09-949-016-129843
;; Sequence 129843; Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CI001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 129843
;; LENGTH: 601

;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-129843

Query Match 83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
|||||
Db 548 CCAAGCTTTTATAT 562

RESULT 13
US-09-949-016-129914
;; Sequence 129914; Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CI001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 129914
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-129914

Query Match 83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
|||||
Db 548 CCAAGCTTTTATAT 562

RESULT 14
US-09-949-016-129985
;; Sequence 129985; Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CI001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 129985
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-129985

Query Match 83.3%; Score 15; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
|||||
Db 548 CCAAGCTTTTATAT 562

RESULT 15
US-09-949-016-130056
; Sequence 130056, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130056
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-130056

Query Match 83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
|||||
Db 548 CCAAGCTTTTATAT 562

Search completed: February 10, 2005, 08:50:04
Job time : 6.53136 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47.34 ; Search time 18.6055 Seconds
(Without alignments)
5727.097 Million cell updates/sec

Title: US-10-790-430-12

Sequence: 1 accaagctttatcatag 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2002as:*
- 6: geneseqn2002bs:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2003ds:*
- 11: geneseqn2004as:*
- 12: geneseqn2004bs:*
- 13: geneseqn2004cs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	6	ABK15249 Corn tran
2	18	100.0	1183	6	ABK15245 Corn nk60
3	16	88.9	33	6	ABL53728 Human BRG
4	16	88.9	107320	11	ACN44736 Mouse gen
5	15	83.3	33	6	ABV74652 Human zin
6	15	83.3	468	4	AAH34926 Human col
7	15	83.3	549	6	ABK76160 Bacillus
8	15	83.3	786	12	ADK17054 Nanoarcha
9	15	83.3	1137	10	ADK30185 Human nov
10	15	83.3	1248	12	ADK30185 Human nov
11	15	83.3	2152	10	AAA48446 Novel mou
12	15	83.3	2152	10	ADC42306 CDNA enco
13	15	83.3	5248	4	ADH61060 Zebrafish
14	15	83.3	8897	4	ABL20304 Drosophi
15	15	83.3	15932	8	ABZ73902 Secreted
16	15	83.3	15932	8	ADA44289 Human sec
17	15	83.3	90885	12	ADK16049 4
18	15	83.3	144792	10	ADC87620 Human GPC
19	15	83.3	144792	10	ADC87620 Human GPC
20	15	83.3	218336	8	ABQ76678 Androgen

ALIGNMENTS

RESULT 1	ABK15249	standard; DNA, 18 BP.
ID	ABK15249	
XX	ABK15249;	
AC	29-AUG-2003 (revised)	
DT	08-MAY-2002 (first entry)	
XX	Corn transgene junction probe #4.	
DE	Probe; ss; rice actin 1 promoter; RAI; RAI intron;	
KW	chloroplast transit peptide gene; glycosylase resistance; corn;	
KW	5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;	
KW	transcriptional terminator; cauliflower mosaic virus 35S promoter;	
KW	PV-ZMG132; transgenic; nk603.	
XX	Agrobacterium tumefaciens.	
OS	Oryza sativa.	
OS	Chimeric.	
XX	EP1167531-A1.	
PN	02-JAN-2002.	
PD	15-JUN-2001; 2001EP-00202314.	
XX	22-JUN-2000; 2000US-0213567P.	
PF	13-OCT-2000; 2000US-0240014P.	
PR	13-OCT-2000; 2000US-0241215P.	
PR	(MONS) MONSANTO TECHNOLOGY LLC.	
PA	Behr CF, Hironaka C, Heck GR, You J;	
PI	WPI; 2002-165871/22.	
XX	Novel DNA construct useful for producing a corn plant that tolerates	
PT	application of glyphosate herbicide, comprises two transgene expression	
PT	cassettes.	
XX	Claim 8; Page 3; 25pp; English.	
PS		
XX		

C	21	14	77.8	337	6	ABN75127 Human ORF
C	22	14	77.8	371	4	AAK67516 Human imm
C	23	14	77.8	377	4	AAK6187 Human imm
C	24	14	77.8	443	9	ACH29826 Human tes
C	25	14	77.8	453	4	AAI11775 Probe #17
C	26	14	77.8	453	4	ABA53465 Human foe
C	27	14	77.8	453	4	AAI33087 Probe #17
C	28	14	77.8	453	4	ABA43053 Human bre
C	29	14	77.8	453	4	ABA23236 Probe #17
C	30	14	77.8	453	4	AAK7192 Human bon
C	31	14	77.8	453	4	AAK01737 Human bra
C	32	14	77.8	453	4	ABS26774 Human liv
C	33	14	77.8	453	5	AAI01706 Probe #16
C	34	14	77.8	453	6	ABS01738 Human gen
C	35	14	77.8	457	3	AAK36726 Arabidops
C	36	14	77.8	470	3	AAK40800 Arabidops
C	37	14	77.8	470	3	AAK45230 Arabidops
C	38	14	77.8	483	4	AAI33833 Probe #25
C	39	14	77.8	483	6	ABS02355 Human gen
C	40	14	77.8	488	6	ABS67082 Human col
C	41	14	77.8	498	6	ABT05259 Human col
C	42	14	77.8	528	10	ABX57529 Arabidops
C	43	14	77.8	636	3	AAK14827 Aspergill
C	44	14	77.8	687	2	AAV74919 Staphyloc
C	45	14	77.8	761	6	ABT09294 Phase-1 R

XX The present invention relates to novel human ERG conjugated protein 72
 CC (see ABB75713). The protein and its coding sequence are useful for
 CC treatment of several diseases, such as medullary leukaemia, lipoma,
 CC inflammation, immunological disease, haemopathy, development disturbance
 CC and HIV infection. The present sequence is a PCR primer, which was used
 CC in an example from the invention
 CC
 XX Sequence 33 BP; 8 A; 6 C; 1 G; 18 T; 0 U; 0 Other;
 SQ
 Query Match 88.9%; Score 16; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CCAAGCTTTATATA 17
 DB 2 CCAAGCTTTATATA 17
 RESULT 4
 ACN44736/c
 ID ACN44736 standard; DNA; 107320 BP.
 AC ACN44736;
 DT 18-NOV-2004 (first entry)
 XX Mouse genomic sequence MCG2943.
 DE
 XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
 XX
 XX Mus musculus.
 OS
 XX WO2003073826-A2.
 PN 12-SEP-2003.
 PD 28-FEB-2003; 2003WO-US006235.
 XX 01-MAR-2002; 2002US-00087192.
 PR
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 XX Morris DW;
 PI
 DR WPI; 2003-328604/31.
 PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 XX
 XX Claim 1; SEQ ID NO 1333; Opp; English.
 XX
 CC The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published
 CC
 XX Sequence 107320 BP; 30076 A; 19214 C; 21901 G; 35227 T; 0 U; 902 Other;
 SQ
 Query Match 88.9%; Score 16; DB 11; Length 107320;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTATATA 17
 DB 79987 CCAAGCTTTATATA 79972
 RESULT 5
 ABV74652
 ID ABV74652 standard; DNA; 33 BP.
 XX
 AC ABV74652;
 DT 03-FEB-2003 (first entry)
 XX
 DE Human zinc finger protein 64.90 PCR primer #4.
 XX
 XX Human; zinc finger protein 64.90; cancer; HIV infection; cytostatic;
 XX anti-HIV; PCR; primer; ss.
 XX
 OS Homo sapiens.
 XX
 EN CN1343696-A.
 XX
 PD 10-APR-2002.
 XX
 PF 19-SEP-2000; 2000CN-00125223.
 XX
 PR 19-SEP-2000; 2000CN-00125223.
 XX
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 DR WPI; 2002-548869/59.
 XX
 XX Novel human zinc finger protein 64.90 for treating cancer and HIV
 XX infection.
 PT
 PT
 XX
 PS Example 5; Page 18 (Disclosure); 34pp; Chinese.
 XX
 CC The present invention relates to Human zinc finger protein 64.90 (see
 CC ABB98772). The zinc finger protein and its coding sequence are useful for
 CC treating cancer and HIV infection. The present sequence is a PCR primer,
 CC which was used in an example from the invention
 CC
 XX Sequence 33 BP; 6 A; 9 C; 4 G; 14 T; 0 U; 0 Other;
 SQ
 Query Match 83.3%; Score 15; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CCAAGCTTTATATA 16
 DB 2 CCAAGCTTTATATA 16
 RESULT 6
 AAH34926
 ID AAH34926 standard; cDNA; 468 BP.
 XX
 AC AAH34926;
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:2008.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 17; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 PD 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.
PF
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI: 2001-235357/24.
DR P-PSDB; AAG75521.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
XX Claim 1; Page 3505; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acid
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
XX to 1052, 7921 and 7922
XX
SQ Sequence 468 BP; 87 A; 140 C; 134 G; 99 T; 0 U; 8 Other;
XX
Query Match 83.3%; Score 15; DB 4; Length 468;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCAAGCTTTATATA 15
Db 454 ACCAAGCTTTATATA 468
XX
RESULT 7
ABK76160/c
ID ABK76160 standard; DNA; 549 BP.
XX
XX ABK76160;
AC
XX
DT 13-AUG-2002 (first entry)
XX
XX Bacillus licheniformis genomic sequence tag (GST) #3451.
DE
XX
XX Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
XX Bacillus licheniformis.
OS
XX
XX WO200229113-A2.
PN
XX
XX 11-APR-2002.
PD
XX
XX 05-OCT-2001; 2001WO-US031437.
PF
XX
XX 06-OCT-2000; 2000US-00680598.
PR
XX 27-MAR-2001; 2001US-0279526P.
XX

PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX
XX
PI Becka R, Clausen IG;
XX
XX WPI: 2002-416684/44.
DR
XX
XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
XX Claim 4; SEQ ID NO 3451; 200pp; English.
XX
XX
XX The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 549 BP; 145 A; 151 C; 118 G; 135 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 6; Length 549;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CCAAGCTTTATATAT 16
Db 504 CCAAGCTTTATATAT 490
XX
RESULT 8
ADK17054
ID ADK17054 standard; DNA; 786 BP.
XX
XX ADK17054;
AC
XX
DT 06-MAY-2004 (first entry)
XX
XX Nanoarchaeum equitans cancer-associated (CA) gene #503.
DE
XX
XX cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia;
KW ds; gene.
XX
XX Nanoarchaeum equitans.
OS
XX
XX WO2003093434-A2.
PN
XX
XX 13-NOV-2003.
PD
XX
XX 01-MAY-2003; 2003WO-US013699.
PF
XX
XX 01-MAY-2002; 2002US-0377447P.
PR
XX
XX (DIVE-) DIVERSA CORP.
PA
XX Stetter KO, Waters E, Kretz K, Podar M, Richardson T;

PI Noordewier M;
XX
XX WPI; 2004-053041/05.
DR P-PSDB; ADK17055.
XX
PT New recombinant cancer-associated genes, such as KCMN9, useful for
PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,
PT cervical, or skin cancers, lymphomas, or leukemia.
XX
PS Claim 5; SEQ ID NO 1006; 251pp; English.
XX
CC The invention comprises then amino acid and coding sequences of cancer-
CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention
CC also comprises the Nanoarchaeum equitans genome. The DNA and protein
CC sequences of the invention are useful for diagnosing and treating cancer
CC (e.g. carcinoma, lymphoma, or leukemia). The present DNA sequence
CC represents a Nanoarchaeum equitans CA gene of the invention.
XX
SQ Sequence 786 BP; 295 A; 112 C; 135 G; 244 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 12; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CAAGCTTTTATATA 17
DB 584 CAAGCTTTTATATA 598
XX
RESULT 9
AD30185/c
ID AD30185 standard; cDNA; 1137 BP.
XX
AC AD30185;
XX
DT 18-DEC-2003 (first entry)
DE Human novel cDNA sequence, SEQ ID NO:267.
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KM ulcers; osteoporosis; autoimmune disease; cancer;
KM molecular weight marker; food supplement; antiparkinsonian; nootropic;
KM neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KM antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KM gene therapy; gene; ss.
XX
XX Homo sapiens.
OS
XX WO2003029271-A2.
PN
XX 10-APR-2003.
PD
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QH, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI; 2003-371981/35.
DR P-PSDB; AD31156.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 267; 1185pp; English.

XX
CC The invention relates to 971 novel human cDNA sequences (AD29919-
CC AD30889) and the polypeptides they encode (AD30890-AD31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (AD31861-AD32627) and the polypeptides encoded by the contigs (AD32628
CC -AD33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1137 BP; 295 A; 260 C; 292 G; 290 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAAGCTTTTATATA 15
DB 398 ACCAAGCTTTTATATA 384
XX
RESULT 10
AD35611/c
ID AD35611 standard; DNA; 1248 BP.
XX
AC AD35611;
XX
DT 26-AUG-2004 (first entry)
DE Novel mouse gene sequence #284.
XX
XX mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KM ischemic heart disease; thrombosis; immune disorder; bacterial disorder;
KM viral disorder; ds; gene.
XX
XX Mus sp.
OS
XX WO2004046310-A2.
PN
XX 03-JUN-2004.
PD
XX 24-OCT-2003; 2003WO-US033948.
PF
XX
XX 15-NOV-2002; 2002US-0426916P.
PR 04-DEC-2002; 2002US-0431158P.
PR 05-DEC-2002; 2002US-0431445P.
PR 05-DEC-2002; 2002US-0431606P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.

PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485359P.
PR 08-AUG-2003; 2003US-0493332P.
PR 08-AUG-2003; 2003US-0493356P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
DR WPI; 2004-431966/40.
XX
PT New mouse nucleic acid molecules and polypeptides, useful for treating
PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
PT disease or thrombosis.
XX
PS Claim 1; SEQ ID NO 284; 263bp; English.
XX
CC The invention comprises 744 novel mouse DNA sequences (genes). The DNA
CC sequences of the invention are useful for treating cancer, psoriasis,
CC ulcerative colitis, inflammation, ischemic heart disease, thrombosis,
CC immune disorders, bacterial disorders and viral disorders. The present
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
CC present DNA sequence is not shown in the specification, but has been
CC retrieved from the WIFO website.
XX
SQ Sequence 1248 BP; 452 A; 245 C; 236 G; 315 T; 0 U; 0 Other;
XX
Query Match 83.3%; Score 15; DB 12; Length 1248;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCAAGCTTTTATAT 16
DB 931 CCAAGCTTTTATAT 917
XX
RESULT 11
AAA48446/c
ID AAA48446 standard; cDNA; 2152 BP.
XX
AC AAA48446;
XX
DT 15-SEP-2003 (revised)
DT 27-OCT-2000 (first entry)
XX
DE Zebrafish parathyroid hormone type-3 receptor PTH3R coding sequence.
XX
KW Zebrafish; parathyroid hormone type-3 receptor; PTH3R;
KW developmental disorder; physiological disorder; neurological disorder;
KW ss.
XX
OS Danio rerio.
XX
FH Key Location/Qualifiers
FT CDS 394..2022
FT /*tag= a
FT /product= "PTH3R"
XX
PN MO200032775-A1.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US028207.
XX
PR 30-NOV-1998; 98US-0110467P.
XX
PA (JUEP/) JUEPPNER H.
PA (RUBI/) RUBIN D A.
XX
PI Jueppner H, Rubin DA;
XX
DR WPI; 2000-412323/35.
DR P-PSDB; AAY99601.

XX
PT New nucleic acids encoding parathyroid hormone receptors PTH3R and PTH3R,
PT useful for treating diseases or disorders associated with impaired
PT receptor functions comprises a specific nucleotide sequence.
XX
PS Claim 23; Fig 1D; 111bp; English.
XX
CC The present sequence is the parathyroid hormone type-3 receptor (PTH3R)
CC coding sequence from the zebrafish. It was obtained by sequencing a cDNA
CC clone. The gene and protein can be used to detect diseases in man where
CC the receptor is either overexpressed or underexpressed, and they can be
CC used to treat these diseases, which may be developmental, physiological
CC or neurological disorders. They can also be used to identify agonists and
CC antagonists which can be used in a similar manner. In addition, the gene
CC can be used for chromosome identification. (Updated on 15-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 2152 BP; 592 A; 471 C; 529 G; 559 T; 0 U; 1 Other;
XX
Query Match 83.3%; Score 15; DB 3; Length 2152;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAAGCTTTTATTA 15
DB 2090 ACCAAGCTTTTATTA 2076
XX
RESULT 12
ADC42306/c
ID ADC42306 standard; cDNA; 2152 BP.
XX
AC ADC42306;
XX
DT 18-DEC-2003 (first entry)
XX
DE cDNA encoding zebrafish parathyroid hormone receptor PTH3R seq id 3.
XX
KW parathyroid hormone; PTH; PTH-related peptide; PTHrP;
KW parathyroid hormone receptor; PTHR; chromosome identification; zebrafish;
KW PTH3R; gene; ss.
XX
OS Danio rerio.
XX
FH Key Location/Qualifiers
FT CDS 394..2022
FT /*tag= a
FT /product= "Zebrafish PTH3R"
FT /note= "Parathyroid hormone receptor"
XX
PN US6541220-B1.
XX
PD 01-APR-2003.
XX
PF 30-NOV-1999; 99US-00449632.
XX
PR 30-NOV-1998; 98US-0110467P.
XX
PA (GEO) GEN HOSPITAL CORP.
XX
PI Jueppner H, Rubin DA;
XX
DR WPI; 2003-754511/71.
DR P-PSDB; ADC42307.
XX
PT Novel nucleic acid comprising a polynucleotide encoding parathyroid
PT hormone/parathyroid hormone-related peptide receptor, useful for
PT chromosome identification.
XX
PS Disclosure; SEQ ID NO 3; 52bp; English.
XX
CC The invention describes an isolated nucleic acid (I) comprising a
CC polynucleotide having a nucleotide sequence chosen from nucleotide

CC sequence encoding a parathyroid hormone (PTH)/PTH-related peptide (PTHrP)
 CC receptor (PTHrP receptor) having a fully defined sequence of 536 amino
 CC acids as given in the specification, PTHrP receptor, mature PTHrP
 CC receptor, PTHrP extracellular or transmembrane domain, and their
 CC complement. (I) is useful for diagnosing and treating decrease in the
 CC standard or normal level of PTHrP receptor activity in an individual, and
 CC for chromosome identification. This sequence encodes zebrafish PTHrP.
 XX

Sequence 2152 BP; 592 A; 470 C; 529 G; 560 T; 0 U; 1 Other;

Query Match 83.3%; Score 15; DB 10; Length 2152;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAGCTTTATTA 15
 |||||
 DB 2090 ACCAGCTTTATTA 2076

RESULT 13
 ADH61060/c
 ID ADH61060 standard; cDNA; 2152 BP.
 XX
 AC ADH61060;
 XX
 DT 25-MAR-2004 (first entry)
 XX
 DE Zebrafish PTHrP cDNA from clone zPTHrP.
 XX
 KW Zebrafish; parathyroid hormone; PTH; parathyroid hormone related protein;
 KM PTHrP; parathyroid hormone related protein receptor; PTHrP; PTHrP;
 KM diagnosis; prognosis; pharmaceutical composition; chromosome assay; gene;
 KM ss.
 XX
 OS Danio rerio.
 XX
 XX Key location/Qualifiers
 FH 394..2022
 FT CDS /*tag= b
 FT /*product= "Zebrafish PTHrP protein"
 FT sig_peptide 394..457
 FT /*tag= a
 FT mat_peptide 458..2019
 FT /*tag= c
 FT /*product= "Zebrafish mature PTHrP protein"
 FT
 XX US2003162256-A1.
 XX
 XX 28-AUG-2003.
 XX
 XX 25-FEB-2003; 2003US-00372095.
 XX
 XX 30-NOV-1998; 98US-0110467P.
 XX
 XX 30-NOV-1999; 99US-0049632.
 XX
 XX (MARS-) MASSACHUSETTS GEN HOSPITAL.
 XX
 XX Jueppner H, Rubin DA;
 PI WPI; 2003-897927/82.
 XX
 XX P-PSDB; ADH61061.
 XX
 XX New parathyroid hormone receptors designated PTHrP and PTHrP isolated
 PT from zebrafish are useful to diagnose and treat parathyroid hormone
 PT receptor-related diseases.
 PT
 XX
 XX Claim 23; SEQ ID NO 3; 53bp; English.
 XX
 XX The present invention relates to novel parathyroid hormone (PTH) and
 CC parathyroid hormone related protein (PTHrP) receptors PTHrP and PTHrP
 CC isolated from zebrafish. The invention is useful in the diagnosis and
 CC prognosis of certain diseases and disorders that express significantly
 CC decreased levels of PTHrP and PTHrP. The invention is also useful in

CC preparing pharmaceutical compositions and in chromosome assays. The
 CC present sequence is zebrafish PTHrP cDNA.
 XX

Sequence 2152 BP; 592 A; 470 C; 529 G; 560 T; 0 U; 1 Other;

Query Match 83.3%; Score 15; DB 10; Length 2152;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAGCTTTATTA 15
 |||||
 DB 2090 ACCAGCTTTATTA 2076

RESULT 14
 ABL10414/c
 ID ABL10414 standard; cDNA; 5248 BP.
 XX
 AC ABL10414;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 25724.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 XX 23-MAR-2000; 2000US-0191637P.
 XX
 XX 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX
 XX P-PSDB; ABB66311.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 PT
 XX
 XX Claim 1; SEQ ID NO 25724; 21bp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL20511), expressed DNA
 CC sequences (ABL101840-ABL16175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences
 CC
 XX Sequence 5248 BP; 1525 A; 1056 C; 1162 G; 1505 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 4; Length 5248;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CAAGCTTTATTA 17
 |||||
 DB 4575 CAAGCTTTATTA 4561

RESULT	15
ID	ABL20304/C
AC	ABL20304 standard; DNA; 8897 BP.
XX	
XX	ABL20304;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 12385.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ds.
XX	
OS	Drosophila melanogaster.
PX	WO200171042-A2.
PD	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US009231.
XX	
PR	23-MAR-2000; 2000US-0191637P.
PR	11-JUL-2000; 2000US-00614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
PT	
PS	Claim 1; SEQ ID NO 12385; 21bp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLA6176-ABLB30511), expressed DNA sequences (ABE01840-ABLA6175) and the encoded proteins (ABB57737-ABB872072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC	
SQ	Sequence 8897 BP; 2656 A; 1738 C; 1855 G; 2648 T; 0 U; 0 Other;
Query Match	83.3%; Score 15; DB 4; Length 8897;
Best Local Similarity	100.0%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	3 CAAGCTTTATATA 17
DB	5922 CAAGCTTTATATA 5908

Search completed: February 10, 2005, 13:44:15
Job time : 22.6055 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 29.935 Seconds
(without alignments)
3460.943 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18

Sequence: 1 accaagctttatcatag 18

Scoring table: OLIGO NUC
Gapop_60.0, Gapext 60.0

Searched: 4313606 seqs, 287787103 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	18	100.0	18	9	US-09-872-051-12
2	18	100.0	18	18	US-10-790-430-12
3	18	100.0	1183	9	US-09-872-051-8
4	18	100.0	1183	18	US-10-790-430-8
5	16	88.9	612	18	US-10-425-115-76005
6	16	88.9	107320	13	US-10-087-192-1333
7	15	83.3	440	17	US-10-424-599-142571
8	15	83.3	468	15	US-10-106-698-2018
9	15	83.3	533	13	US-10-027-632-194059
10	15	83.3	533	17	US-10-027-632-194059
11	15	83.3	549	9	US-09-974-300-3451

C 12	15	83.3	613	13	US-10-027-632-194858	Sequence 194858,
C 13	15	83.3	613	13	US-10-027-632-194859	Sequence 194859,
C 14	15	83.3	613	17	US-10-027-632-194858	Sequence 194858,
C 15	15	83.3	613	17	US-10-027-632-194859	Sequence 194859,
C 16	15	83.3	996	13	US-10-027-632-31469	Sequence 31469, A
C 17	15	83.3	996	17	US-10-027-632-31469	Sequence 31469, A
C 18	15	83.3	2152	16	US-10-372-095-3	Sequence 3, Appl1
C 19	15	83.3	9224	13	US-10-108-605-254	Sequence 254, Appl
C 20	15	83.3	94463	18	US-10-719-993-6944	Sequence 6944, Ap
C 21	15	83.3	218356	18	US-10-473-933-10	Sequence 10, Appl
C 22	15	83.3	561515	18	US-10-741-601-5682	Sequence 5682, Ap
C 23	15	83.3	744802	17	US-10-292-798-1369	Sequence 1369, Ap
C 24	14	77.8	194	18	US-10-425-115-69101	Sequence 69101, A
C 25	14	77.8	201	18	US-10-719-993-24777	Sequence 24777, A
C 26	14	77.8	201	18	US-10-719-993-25079	Sequence 25079, A
C 27	14	77.8	201	18	US-10-719-993-25081	Sequence 25081, A
C 28	14	77.8	206	18	US-10-767-701-16656	Sequence 16656, A
C 29	14	77.8	290	18	US-10-674-124A-12783	Sequence 12783, A
C 30	14	77.8	337	11	US-09-864-408A-147	Sequence 147, App
C 31	14	77.8	430	17	US-10-424-599-96924	Sequence 96924, A
C 32	14	77.8	431	18	US-10-425-115-10849	Sequence 10849, A
C 33	14	77.8	432	18	US-10-425-115-18546	Sequence 18546, A
C 34	14	77.8	443	10	US-09-918-995-17038	Sequence 17038, A
C 35	14	77.8	453	9	US-09-864-761-1702	Sequence 1702, Ap
C 36	14	77.8	469	17	US-10-242-535A-39839	Sequence 39839, A
C 37	14	77.8	469	17	US-10-085-783A-39839	Sequence 39839, A
C 38	14	77.8	482	13	US-10-027-632-77058	Sequence 77058, A
C 39	14	77.8	482	17	US-10-027-632-77058	Sequence 77058, A
C 40	14	77.8	483	9	US-09-864-761-15448	Sequence 15448, A
C 41	14	77.8	489	17	US-10-424-599-37844	Sequence 37844, A
C 42	14	77.8	498	9	US-09-989-919-43	Sequence 43, Appl
C 43	14	77.8	528	9	US-09-770-152-881	Sequence 881, App
C 44	14	77.8	611	13	US-10-027-632-213307	Sequence 213307,
C 45	14	77.8	611	13	US-10-027-632-213308	Sequence 213308,

ALIGNMENTS

RESULT 1
US-09-872-051-12
Sequence 12, Application US/09872051
Patent No. US20020013960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinong
TITLE OF INVENTION: Corn Event PV-ZMGT32(mk603) and Composition and Methods for Der
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(18)
OTHER INFORMATION: Agrobacterium tumefaciens nos 3' terminator and rice actin pro
US-09-872-051-12

Query Match 100.0%; Score 18; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18
|||
1 ACCAAGCTTTATATAG 18

RESULT 2

US-10-790-430-12
; Sequence 12, Application US/10790430
; Publication No. US20040139493A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detect
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/10/790,430
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(18)
; OTHER INFORMATION: Agrobacterium tumefaciens nos 3' terminator and rice actin promot
; OTHER INFORMATION: er DN
US-10-790-430-12

Query Match 100.0%; Score 18; DB 18; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18
|||
1 ACCAAGCTTTATATAG 18

RESULT 3

US-09-872-051-8
; Sequence 8, Application US/09872051
; Patent No. US20020013960A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detect
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

Query Match 100.0%; Score 18; DB 9; Length 1183;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18
|||
156 ACCAAGCTTTATATAG 173

RESULT 4

US-10-790-430-8
; Sequence 8, Application US/10790430
; Publication No. US20040139493A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detect
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/10/790,430
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-10-790-430-8

Query Match 100.0%; Score 18; DB 18; Length 1183;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18
|||
156 ACCAAGCTTTATATAG 173

```
RESULT 5
US-10-425-115-76005/c
; Sequence 76005, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 76005
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169344C.1
US-10-425-115-76005

Query Match      88.9%; Score 16; DB 18; Length 612;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 CAAAGCTTTATATAG 18
DB      146 CAAAGCTTTATATAG 131

RESULT 6
US-10-087-192-1333/c
; Sequence 1333, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1333
; LENGTH: 107320
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) .. (107320)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1333

Query Match      88.9%; Score 16; DB 13; Length 107320;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCAAGCTTTATATA 17
DB      79987 CCAAGCTTTATATA 79972

RESULT 7
US-10-424-599-142571/c
; Sequence 142571, Application US/10424599
; Publication No. US20040031072A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142571
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99757C.1
US-10-424-599-142571

Query Match      83.3%; Score 15; DB 17; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAGCTTTATATATAG 18
DB      264 AAGCTTTATATATAG 250

RESULT 8
US-10-106-698-2018
; Sequence 2018, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2018
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (434) .. (434)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2018

Query Match      83.3%; Score 15; DB 15; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAAGCTTTATATA 15
DB      454 ACCAAGCTTTATATA 468

RESULT 9
US-10-027-632-194059
; Sequence 194059, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
```

FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194059
LENGTH: 533
TYPE: DNA
ORGANISM: Human
US-10-027-632-194059

Query Match 83.3%; Score 15; DB 13; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
DB 390 CCAAGCTTTTATAT 404

RESULT 10
US-10-027-632-194059
Sequence 194059, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194059
LENGTH: 533
TYPE: DNA
ORGANISM: Human
US-10-027-632-194059

Query Match 83.3%; Score 15; DB 17; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
DB 390 CCAAGCTTTTATAT 404

DB 390 CCAAGCTTTTATAT 404
RESULT 11
US-09-974-300-3451/C
Sequence 3451, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
TITLE OF INVENTION: Expression
FILE REFERENCE: 10085,500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3451
LENGTH: 549
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-3451

Query Match 83.3%; Score 15; DB 9; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
DB 504 CCAAGCTTTTATAT 490

RESULT 12
US-10-027-632-194858/C
Sequence 194858, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194858
LENGTH: 613
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc feature
LOCATION: (1)_(613)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-194858

Query Match 83.3%; Score 15; DB 9; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16
DB 504 CCAAGCTTTTATAT 490

Query Match 83.3% Score 15; DB 13; Length 613;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17
|||
Db 142 CAAGCTTTTATATA 128

RESULT 13

US-10-027-632-194859/c
; Sequence 194859, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194859
; LENGTH: 613
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(613)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-194859

Query Match 83.3% Score 15; DB 13; Length 613;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17
|||
Db 142 CAAGCTTTTATATA 128

RESULT 14

US-10-027-632-194858/c
; Sequence 194858, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194858
; LENGTH: 613
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(613)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-194858

Query Match 83.3% Score 15; DB 17; Length 613;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17
|||
Db 142 CAAGCTTTTATATA 128

RESULT 15

US-10-027-632-194859/c
; Sequence 194859, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194859
; LENGTH: 613
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(613)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-194859

Query Match 83.3% Score 15; DB 17; Length 613;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17
|||
Db 142 CAAGCTTTTATATA 128

Mon Feb 14 10:12:59 2005

us-10-790-430-12.oligo.rmpb

Page 6

Search completed: February 9, 2005, 11:12:47
Job time : 30.935 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 121.68 Seconds
(without alignments)
5630.828 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18
Sequence: 1 accaagctttatataag 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17	94.4	651	9	CE238229
C 2	16	88.9	118	5	BP008680
C 3	16	88.9	196	5	BM295435
C 4	16	88.9	199	1	AV906194
C 5	16	88.9	280	9	CI225702
C 6	16	88.9	298	1	AV906393
C 7	16	88.9	357	8	B39844
C 8	16	88.9	363	1	AV891743
C 9	16	88.9	382	8	B55409
C 10	16	88.9	431	8	B45104
C 11	16	88.9	432	8	B98924
C 12	16	88.9	476	8	AQ829617
C 13	16	88.9	527	1	AV903514
C 14	16	88.9	572	1	AV904180
C 15	16	88.9	618	5	BP016541
C 16	16	88.9	627	5	BP004638
C 17	16	88.9	635	1	AV895582
C 18	16	88.9	640	8	B20453
C 19	16	88.9	651	5	BM167398
C 20	16	88.9	655	1	AV897075
C 21	16	88.9	657	5	BM065072
C 22	16	88.9	659	5	BM192868
C 23	16	88.9	665	5	BM065061
C 24	16	88.9	666	1	AV906223

C 25	16	88.9	667	5	BM066300
C 26	16	88.9	668	1	AV893354
C 27	16	88.9	668	5	BM191586
C 28	16	88.9	669	1	AV893491
C 29	16	88.9	670	5	BP016388
C 30	16	88.9	672	5	BM191597
C 31	16	88.9	674	5	BP002271
C 32	16	88.9	678	5	BP008293
C 33	16	88.9	682	5	BP006864
C 34	16	88.9	683	5	BM296842
C 35	16	88.9	695	5	BP004801
C 36	16	88.9	717	6	CA174264
C 37	16	88.9	834	8	BZ466621
C 38	16	88.9	862	5	BU459296
C 39	16	88.9	928	6	CA243532
C 40	16	88.9	1358	9	AG463618
C 41	15	83.3	128	8	CC380153
C 42	15	83.3	158	8	B26837
C 43	15	83.3	171	9	AL943005
C 44	15	83.3	187	7	Z38858
C 45	15	83.3	193	4	BI490751

ALIGNMENTS

RESULT 1
CE238229/c 651 bp DNA linear GSS 25-SEP-2003
LOCUS
DEFINITION
tigr-gss-dog-1700033406145 Dog Library Canis familiaris genomic,
genomic survey sequence.

ACCESSION
CE238229 GI:35393987
VERSION
CE238229.1
KEYWORDS
GSS.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
1 (bases 1 to 651)
Kirkness,E.F., Batina,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: Shotgun.

FEATURES
source
location/Qualifiers
1..651
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Bactx; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 651;
Best Local Similarity 100.0%; Pred.No.29;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCAAGCTTTTATATAG 18
DB 606 CCAAGCTTTATATAG 590

```

RESULT 2
LOCUS BP008680 118 bp mRNA linear EST 15-MAR-2002
DEFINITION BP008680 Nori Satoh unpublished cDNA library, young adult Ciona
intestinalis cDNA clone ciad49h02 5', mRNA sequence.
ACCESSION BP008680
VERSION BP008680.1 GI:19500157
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 118)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
location/Qualifiers
1..118
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciad49h02"
/rissue_type="whole animal"
/dev_stage="young adult"
/clone_id="Nori Satoh unpublished cDNA library, young
adult"

ORIGIN
Query Match 88.9%; Score 16; DB 5; Length 118;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATAT 16
|||||
21 ACCAAGCTTTATAT 36

RESULT 3
LOCUS BM295435 196 bp mRNA linear EST 11-NOV-2002
DEFINITION BM295435 Nori Satoh unpublished cDNA library, neural complex Ciona
intestinalis cDNA clone cinc010p13 5', mRNA sequence.
ACCESSION BM295435
VERSION BM295435.1 GI:24876046
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 196)
AUTHORS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
location/Qualifiers
1..196
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone_id="Nori Satoh unpublished cDNA library, young
adult"

FEATURES
source

```

```

/cione="cinc010p13"
/rissue_type="neural complex"
/clone_id="Nori Satoh unpublished cDNA library, neural
complex"

ORIGIN
Query Match 88.9%; Score 16; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 16+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATAT 16
|||||
54 ACCAAGCTTTATAT 39

RESULT 5
LOCUS CL225702 280 bp DNA linear GSS 08-JAN-2004
DEFINITION ZMMBc0564122r ZMMBc Zea mays genomic clone ZMMBc0564122 3',
genomic survey sequence.
ACCESSION CL225702
VERSION CL225702.1 GI:40761719
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 280)

ORIGIN
Query Match 88.9%; Score 16; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 16+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATAT 16
|||||
54 ACCAAGCTTTATAT 39

RESULT 4
LOCUS AV906194/c 199 bp mRNA linear EST 09-NOV-2001
DEFINITION AV906194 Nori Satoh unpublished cDNA library, young adult Ciona
intestinalis cDNA clone rciad65c15 3', mRNA sequence.
ACCESSION AV906194
VERSION AV906194.1 GI:16895292
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 199)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
location/Qualifiers
1..199
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rciad65c15"
/rissue_type="whole animal"
/dev_stage="young adult"
/clone_id="Nori Satoh unpublished cDNA library, young
adult"

FEATURES
source

```

AUTHORS Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
Zohovetz, V., Puka, G., Yu, Y., Wang, R. and Messing, J.
TITLE Sequencing of the maize genome at PGR (2003c)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..280
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/clone="ZMBC056412"
/lab_host="E. coli DH10B"
/clone_1lb="ZMBC"
/note="Vector: PTARAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN

Query Match 88.9%; Score 16; DB 9; Length 280;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTATATA 17
|||||
51 CCAAGCTTTATATA 66

RESULT 6
AV906393/c 298 bp mRNA linear EST 09-NOV-2001
LOCUS AV906393 Nori Satoh unpublished cDNA library, young adult Ciona
DEFINITION intestinalis cDNA clone rcia465n15 3', mRNA sequence.

ACCESSION AV906393
VERSION AV906393.1 GI:16895491
KEYWORDS EST.

SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Plebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 298)

AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin-1, T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
Location/Qualifiers
1..298
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcia465n15"
/tissue_type="whole animal"
/dev_stage="young adult"
/clone_1lb="Nori Satoh unpublished cDNA library, young adult"

ORIGIN

Query Match 88.9%; Score 16; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATAT 16
|||||
80 ACCAAGCTTTATAT 65

RESULT 7
B39844 357 bp DNA linear GSS 18-OCT-1997
LOCUS B39844
DEFINITION HS-1050-B1-D08-MR-ab1 CIT Human Genomic Sperm Library C Homo
sapiens genomic clone plate=CT 772 Col=15 Row=H, genomic survey
sequence.

ACCESSION B39844
VERSION B39844.1 GI:2544096
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 357)
Tsalicoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.E.
Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
Unpublished (1997)

AUTHORS Mahairas, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schmidt, S.,
Tralcoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.E.

TITLE Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors

JOURNAL Unpublished (1997)
COMMENT Contact: Mahairas, G.G., Zackrone, K.D., Hood, L.
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackron@u.washington.edu

Sequence Tagged Connector
Plate: CT 772 row: H column: 15
Class: BAC ends
High quality sequence stop: 357.

FEATURES
source
Location/Qualifiers
1..357
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=CT 772 Col=15 Row=H"
/sex="M"
/clone_1lb="CIT Human Genomic Sperm Library C"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 357;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CAAGCTTTATATAG 18
|||||
138 CAAGCTTTATATAG 153

RESULT 8
AV891743/c 363 bp mRNA linear EST 09-NOV-2001
LOCUS AV891743 Nori Satoh unpublished cDNA library, young adult Ciona
DEFINITION intestinalis cDNA clone rcia26f24 3', mRNA sequence.

ACCESSION AV891743
VERSION AV891743.1 GI:16880769
KEYWORDS EST.

SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Plebobranchia; Cionidae; Ciona.

REFERENCE

1 (bases 1 to 363)
Satoh, N., Satou, Y., Kohara, Y. and Shin-1, T.

AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin-1, T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT

Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

Location/Qualifiers
1..363
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rciad26124"
/tissue_type="whole animal"
/dev_stage="young adult"
/clone_lib="Nori Satoh unpublished cDNA library, young adult"

ORIGIN

Query Match 88.9%; Score 16; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGCTTTATAT 16
|||||
DB 53 ACCAGCTTTATAT 38

RESULT 9

B55409 382 bp DNA linear GSS 20-JUN-1998
LOCUS CIT-HSP-387A19.TF CIT-HSP Homo sapiens genomic clone 387A19,
DEFINITION genomic survey sequence.
ACCESSION B55409
VERSION B55409.1 GI:2609743
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 382)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Sub,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends

JOURNAL COMMENT
TITLE Unpublished (1997)
JOURNAL COMMENT
TITLE Unpublished (1997)
JOURNAL COMMENT
TITLE Unpublished (1997)

FEATURES

source

Location/Qualifiers
1..382
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:5379336"
/db_xref="taxon:9606"
/clone="387A19"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 382;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CAAGCTTTATATAG 18
|||||
DB 1 CAAGCTTTATATAG 16

RESULT 10

B45104 431 bp DNA linear GSS 21-OCT-1997
LOCUS HS-1060-B1-F03-NR.abi CIT Human Genomic Sperm Library C Homo
DEFINITION sapiens genomic clone Plate=CT 782 Col=5 Row=L, genomic survey
sequence.
ACCESSION B45104
VERSION B45104.1 GI:2549938
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 431)
Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
Unpublished (1997)
Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackrone@u.washington.edu
Sequence Tagged Connector
Plate: CT 782 row: L column: 5
Class: BAC ends
High quality sequence stop: 431.
Location/Qualifiers
1..431
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=CT 782 Col=5 Row=L"
/sex="M"
/clone_lib="CIT Human Genomic Sperm Library C"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

FEATURES

source

Location/Qualifiers
1..431
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=CT 782 Col=5 Row=L"
/sex="M"
/clone_lib="CIT Human Genomic Sperm Library C"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CAAGCTTTATATAG 18
|||||
DB 50 CAAGCTTTATATAG 65

RESULT 11

B98924 432 bp DNA linear GSS 26-JUN-1998
LOCUS CIT-HSP-2174P18.TF CIT-HSP Homo sapiens genomic clone 2174P18,
DEFINITION genomic survey sequence.
ACCESSION B98924
VERSION B98924.1 GI:3026734
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 432)

REFERENCE

1 (bases 1 to 432)

AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Base,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)

JOURNAL
Unpublished (1998)

COMMENT
Other_GSSs: CIT-HSP-2174P18.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

FEATURES
source
1..432
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7104820"
/db_xref="taxon:9606"
/clone="2174P18"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 88.9%; Score 16; DB 8; Length 432;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
3 CAAGCTTTTATATAG 18
|||||
29 CAAGCTTTTATATAG 44

Db
29 CAAGCTTTTATATAG 44

RESULT 12
AO829617 476 bp DNA linear GSS 27-AUG-1999
AO829617
HS 4832_A1_B01_SP6B CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=4832 Col=1 Row=1, genomic survey sequence.

ACCESSION
AO829617
AO829617 GI:5795679

VERSION
GSS.

KEYWORDS
Homo sapiens (human)

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 476)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

REFERENCE
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu

COMMENT
PUBMED
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu

Plate: 4832 row: 1 column: 1
Seq primer: r7
Class: BAC ends
High quality sequence stop: 476.
Location/Qualifiers
1..476
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=4832 Col=1 Row=1"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

ORIGIN
Query Match 88.9%; Score 16; DB 8; Length 476;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
3 CAAGCTTTTATATAG 18
|||||
177 CAAGCTTTTATATAG 192

Db
177 CAAGCTTTTATATAG 192

RESULT 13
AV903514/c 527 bp mRNA linear EST 09-NOV-2001
AV903514
LOCUS
AV903514 Nori Satoh unpublished cDNA library, young adult Ciona intestinalis cDNA clone rcid460k21 3', mRNA sequence.
AV903514
AV903514.1 GI:16892612

ACCESSION
AV903514.1 GI:16892612

VERSION
EST.

KEYWORDS
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Plebobranchia; Cionidae; Ciona.

ORGANISM
Ciona intestinalis

SOURCE
Ciona intestinalis

REFERENCE
1 (bases 1 to 527)
Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
Expressed genes in Ciona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Kyoto-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

COMMENT
Nori Satoh unpublished cDNA library, young adult

QY
1 ACCAAGCTTTATAT 16
|||||
80 ACCAAGCTTTATAT 65

Db
80 ACCAAGCTTTATAT 65

RESULT 14
AV904180 572 bp mRNA linear EST 09-NOV-2001
AV904180/c
LOCUS
AV904180 Nori Satoh unpublished cDNA library, young adult Ciona

DEFINITION
AV904180 Nori Satoh unpublished cDNA library, young adult Ciona

ACCESSION AV904180
 VERSION GI:16893278
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Plebobranchia; Clonidae; Ciona.
 1 (bases 1 to 572)
 Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
 Expressed genes in Ciona intestinalis
 Unpublished (2000)
 JOURNAL Contact: Nori Satoh
 COMMENT Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 location/Qualifiers
 1. 572
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rciad62115"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young
 adult"

ORIGIN

Query Match 88.9%; Score 16; DB 1; Length 572;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCAAGCTTTATAT 16
 |||||||||
 DB 42 ACCAAGCTTTATAT 27

RESULT 15 618 bp mRNA linear EST 15-MAR-2002
 BP016541
 LOCUS BP016541 Nori Satoh unpublished cDNA library, young adult Ciona
 DEFINITION intestinalis cDNA clone ciad65n15 5', mRNA sequence.
 ACCESSION BP016541
 VERSION BP016541.1 GI:19508018
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Plebobranchia; Clonidae; Ciona.
 1 (bases 1 to 618)
 Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
 Expressed genes in Ciona intestinalis
 Unpublished (2000)
 JOURNAL Contact: Nori Satoh
 COMMENT Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 location/Qualifiers
 1. 618
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="ciad65n15"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young
 adult"

FEATURES
 source
 1. 618
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="ciad65n15"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young
 adult"

ORIGIN

Query Match 88.9%; Score 16; DB 5; Length 618;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCAAGCTTTATAT 16
 |||||||||
 DB 592 ACCAAGCTTTATAT 607
 Search completed: February 10, 2005, 17:02:02
 Job time : 124.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41; Search time 828.202 Seconds
(without alignments)
3460.943 Million cell updates/sec

Title: US-10-790-430-7

Perfect score: 498
Sequence: 1 aatcgatcccaaatcgcgcac.....aaaagtcgccaagtga 498

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	100.0	498	9 US-09-872-051-7	Sequence 7, Appl1
2	498	100.0	498	18 US-10-790-430-7	Sequence 8, Appl1
3	118	23.7	1183	9 US-09-872-051-8	Sequence 8, Appl1
4	118	23.7	1183	18 US-10-790-430-8	Sequence 27, Appl1
5	101	20.3	2378	15 US-10-213-791-27	Sequence 49, Appl1
6	70	14.1	1501	14 US-10-012-070A-49	Sequence 44, Appl1
7	70	14.1	1501	14 US-10-012-013-44	Sequence 43, Appl1
8	68	13.7	1259	10 US-09-991-209-43	Sequence 1, Appl1
9	68	13.7	2480	19 US-10-678-588A-1	Sequence 36, Appl1
10	68	13.7	5164	10 US-09-991-209-36	Sequence 25, Appl1
11	68	13.7	5277	10 US-09-991-209-25	Sequence 25, Appl1

12	68	13.7	5295	10	US-09-991-209-38	Sequence 38, Appl1
13	68	13.7	5327	10	US-09-991-209-27	Sequence 27, Appl1
14	68	13.7	5337	10	US-09-991-209-19	Sequence 19, Appl1
15	68	13.7	5337	10	US-09-991-209-23	Sequence 23, Appl1
16	68	13.7	5337	10	US-09-991-209-31	Sequence 31, Appl1
17	68	13.7	5338	10	US-09-991-209-15	Sequence 15, Appl1
18	68	13.7	5338	10	US-09-991-209-17	Sequence 29, Appl1
19	68	13.7	5345	10	US-09-991-209-49	Sequence 17, Appl1
20	68	13.7	5387	10	US-09-991-209-41	Sequence 41, Appl1
21	68	13.7	5395	10	US-09-991-209-21	Sequence 21, Appl1
22	67	13.5	4032	9	US-09-970-921-5	Sequence 5, Appl1
23	63	12.7	399	9	US-09-682-597-5	Sequence 5, Appl1
24	63	12.7	399	18	US-10-727-423-5	Sequence 5, Appl1
25	63	12.7	399	18	US-10-737-487-5	Sequence 5, Appl1
26	62	12.4	6865	10	US-09-845-064-13	Sequence 13, Appl1
27	62	12.4	10003	10	US-09-845-064-21	Sequence 21, Appl1
28	30	6.0	30	9	US-09-682-597-2	Sequence 2, Appl1
29	30	6.0	30	18	US-10-727-423-2	Sequence 2, Appl1
30	30	6.0	30	18	US-10-737-487-2	Sequence 2, Appl1
31	29	5.8	29	9	US-09-872-051-4	Sequence 4, Appl1
32	29	5.8	29	18	US-10-790-430-4	Sequence 4, Appl1
33	22	4.4	22	9	US-09-872-051-13	Sequence 13, Appl1
34	22	4.4	22	9	US-09-872-051-14	Sequence 14, Appl1
35	22	4.4	22	18	US-10-790-430-13	Sequence 13, Appl1
36	22	4.4	22	18	US-10-790-430-14	Sequence 14, Appl1
37	21	4.2	2000	17	US-10-260-238-2562	Sequence 2562, Appl1
38	20	4.0	2107	15	US-10-213-791-29	Sequence 29, Appl1
39	20	4.0	2436	15	US-10-213-791-31	Sequence 31, Appl1
40	20	4.0	3039	15	US-10-232-665-19	Sequence 19, Appl1
41	20	4.0	3039	15	US-10-232-665-21	Sequence 19, Appl1
42	20	4.0	3044	15	US-10-232-665-38	Sequence 38, Appl1
43	20	4.0	3450	15	US-10-232-665-17	Sequence 17, Appl1
44	20	4.0	3455	15	US-10-232-665-16	Sequence 36, Appl1
45	20	4.0	3469	15	US-10-232-665-23	Sequence 23, Appl1

ALIGNMENTS

RESULT 1
US-09-872-051-7
Sequence 7, Appl1
Patent No. US20020013960A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMGR32 (nke603) and Composition and Methods for Det
FILE REFERENCE: 38-21 (52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
PRIOR APPLICATION NUMBER: 2001-06-01
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 498
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
NAME/KEY: source
LOCATION: (1)..(498)
OTHER INFORMATION: 1-304 Zea maize genomic DNA
OTHER INFORMATION: 305-349 construct vector DNA
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA
US-09-872-051-7

SEQ ID NO 8
LENGTH: 1183
TYPE: DNA


```

; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

Query Match      23.7%; Score 118; DB 9; Length 1183;
Best Local Similarity 100.0%; Pred. No. 9.6e-54;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGGTTAACAGCTTACTCGAGGTCA 367
    |||||||
DB 381 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGGTTAACAGCTTACTCGAGGTCA 322

QY 368 TTCATATGCTTGAGAGAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 425
    |||||||
DB 321 TTCATATGCTTGAGAGAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 264

RESULT 4
US-10-790-430-8/c
; Sequence 8, Application US/10790430
; Publication No. US20040139493A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting and Compositions
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/10/790,430
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-10-790-430-8

Query Match      23.7%; Score 118; DB 18; Length 1183;
Best Local Similarity 100.0%; Pred. No. 9.6e-54;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGGTTAACAGCTTACTCGAGGTCA 367
    |||||||
DB 381 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGGTTAACAGCTTACTCGAGGTCA 322

QY 368 TTCATATGCTTGAGAGAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 425
    |||||||
DB 321 TTCATATGCTTGAGAGAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 264
```

```

RESULT 5
US-10-213-791-27
; Sequence 27, Application US/10213791
; Publication No. US20030106096A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/10/213,791
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US/09/441,340
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/108,763
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:expression
; OTHER INFORMATION: cassette comprising a plant promoter linked to an
; OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
; OTHER INFORMATION: transferase, and a termination sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (28)..(965)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (966)..(1423)
; FEATURE:
; NAME/KEY: transite_peptide
; LOCATION: (1440)..(1667)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1668)..(2099)
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (2114)..(2369)
US-10-213-791-27

Query Match      20.3%; Score 101; DB 15; Length 2378;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 GATATCCCTAGGCGCGCGGTTAACAGCTTACTCGAGGTATTCATATGCTGAGAG 384
    |||||||
DB 1 GATATCCCTAGGCGCGCGGTTAACAGCTTACTCGAGGTATTCATATGCTGAGAG 60

QY 385 AGAGTCGGGATAGTCCAAATAAACAAGTAAGATTACC 425
    |||||||
DB 61 AGAGTCGGGATAGTCCAAATAAACAAGTAAGATTACC 101

RESULT 6
US-10-012-070A-49
; Sequence 49, Application US/10012070A
; Publication No. US20030077801A1
; GENERAL INFORMATION:
; APPLICANT: Hawkes, Timothy
; APPLICANT: Warner, Simon
; APPLICANT: Andrews, Christopher
; APPLICANT: Bachoo, Satvinder
; APPLICANT: Pickerill, Andrew
; TITLE OF INVENTION: Herbicide Resistant Plants
; FILE REFERENCE: 50490/UST
; CURRENT APPLICATION NUMBER: US/10/012,070A
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT/GB00/01573
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 57
```

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sp.
US-10-012-070A-49

Query Match 14.1%; Score 70; DB 14; Length 1501;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGG 415
|||||
DB 750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGG 809

QY 416 TAAGATTACC 425
|||||
DB 810 TAAGATTACC 819

RESULT 7

US-10-012-013-44
; Sequence 44; Application US/10012013
; Publication No. US20030079246A1
; GENERAL INFORMATION:

APPLICANT: Hawkes, Timothy
APPLICANT: Warner, Simon
APPLICANT: Andrews, Christopher
APPLICANT: Bachoo, Savinder
APPLICANT: Pickertill, Andrew
TITLE OF INVENTION: Herbicide Resistant Plants
FILE REFERENCE: 50450/UST
CURRENT APPLICATION NUMBER: US/10/012,013
PRIOR FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: PCT/GB00/01572
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sp.
US-10-012-013-44

Query Match 14.1%; Score 70; DB 14; Length 1501;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGG 415
|||||
DB 750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGG 809

QY 416 TAAGATTACC 425
|||||
DB 810 TAAGATTACC 819

RESULT 8

US-09-991-209-43
; Sequence 43; Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:

APPLICANT: Dunn-Coleman, Nigel
APPLICANT: Langdon, Timothy
APPLICANT: Morse, Phillip
TITLE OF INVENTION: Manipulation of the Phenolic Acid
TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
FILE REFERENCE: GC648-2
CURRENT APPLICATION NUMBER: US/09/991,209
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 60/249,608
PRIOR FILING DATE: 2000-11-17

NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1259
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: actin promoter
US-09-991-209-43

Query Match 13.7%; Score 68; DB 10; Length 1259;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGGTA 417
|||||
DB 16 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGGTA 75

QY 418 AGATTACC 425
|||||
DB 76 AGATTACC 83

QY 418 AGATTACC 425
|||||
DB 76 AGATTACC 83

RESULT 9

US-10-678-588A-1/c
; Sequence 1; Application US/10678588A
; Publication No. US2005002266A1
; GENERAL INFORMATION:

APPLICANT: Wu, Jingrui
TITLE OF INVENTION: Water-Deficit-Tolerant Transgenic Plants
FILE REFERENCE: 38-21(52578)C
CURRENT APPLICATION NUMBER: US/10/678,588A
CURRENT FILING DATE: 2003-10-02
PRIOR APPLICATION NUMBER: US 60/415,758
PRIOR FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: US 60/425,157
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: US 60/463,787
PRIOR FILING DATE: 2003-04-11
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2480
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: transcriptional unit comprising promoter, coding sequence for
US-10-678-588A-1

Query Match 13.7%; Score 68; DB 19; Length 2480;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGGTA 417
|||||
DB 2480 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATATAAACAAGGTA 2421

QY 418 AGATTACC 425
|||||
DB 2420 AGATTACC 2413

QY 418 AGATTACC 425
|||||
DB 2420 AGATTACC 2413

RESULT 10

US-09-991-209-36
; Sequence 36; Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:

APPLICANT: Dunn-Coleman, Nigel
APPLICANT: Langdon, Timothy
APPLICANT: Morse, Phillip
TITLE OF INVENTION: Manipulation of the Phenolic Acid
TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted

```

; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 5164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pP010.1 vector
US-09-991-209-36

Query Match      13.7%; Score 68; DB 10; Length 5164;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 417
          |||||
DB      3934 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 3993

QY      418 AGATTACC 425
          |||||
DB      3994 AGATTACC 4001

RESULT 11
US-09-991-209-25
; Sequence 25, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 5277
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTP5-1 vector
US-09-991-209-25

Query Match      13.7%; Score 68; DB 10; Length 5277;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 417
          |||||
DB      4050 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 4109

QY      418 AGATTACC 425
          |||||
DB      4110 AGATTACC 4117

RESULT 12
US-09-991-209-38
; Sequence 38, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 5295
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pUG4 vector
US-09-991-209-38

Query Match      13.7%; Score 68; DB 10; Length 5295;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 417
          |||||
DB      4068 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 4127

QY      418 AGATTACC 425
          |||||
DB      4128 AGATTACC 4135

RESULT 13
US-09-991-209-27
; Sequence 27, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 5327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTP4a2 vector
US-09-991-209-27

Query Match      13.7%; Score 68; DB 10; Length 5327;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 417
          |||||
DB      3876 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATGTCCTCAAAATTAACCAAGGTA 3935

QY      418 AGATTACC 425
          |||||
DB      3936 AGATTACC 3943
```

RESULT 14
US-09-991-209-19
; Sequence 19, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 5337
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PTU4 vector
US-09-991-209-19

Query Match 13.7%; Score 68; DB 10; Length 5337;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 417
DB 4110 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 4169

QY 418 AGATTACC 425
DB 4170 AGATTACC 4177

RESULT 15
US-09-991-209-23
; Sequence 23, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; APPLICANT: Morse, Phillip
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 5337
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PTP8-5 vector
US-09-991-209-23

Query Match 13.7%; Score 68; DB 10; Length 5337;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 417
DB 4110 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 4169

QY 418 AGATTACC 425
DB 4170 AGATTACC 4177

Search completed: February 9, 2005, 11:12:26
Job time : 829.202 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 5.53136 Seconds
(without alignments)
5324.730 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 1 TGCCTCTCTGCTGACTTC 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patente NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	18	4	US-09-872-051-11	Sequence 11, Appl
2	100.0	183	4	US-09-872-051-8	Sequence 8, Appl
3	88.9	601	4	US-09-949-016-81715	Sequence 81715, A
4	88.9	601	4	US-09-949-016-81717	Sequence 81717, A
5	88.9	31391	4	US-09-949-016-14319	Sequence 14319, A
6	88.9	125192	4	US-09-949-016-14120	Sequence 14120, A
7	83.3	21	4	US-09-422-978-7605	Sequence 7605, Ap
8	83.3	601	4	US-09-949-016-58184	Sequence 58184, A
9	83.3	21125	4	US-09-949-016-15108	Sequence 15108, A
10	83.3	57761	4	US-09-949-016-13429	Sequence 13429, A
11	77.8	20	4	US-09-198-452A-5457	Sequence 5457, Ap
12	77.8	241	4	US-09-016-434-403	Sequence 403, App
13	77.8	356	4	US-09-621-976-9793	Sequence 9793, Ap
14	77.8	459	4	US-09-489-039A-4	Sequence 4, Appl
15	77.8	601	4	US-09-949-016-19121	Sequence 19121, A
16	77.8	601	4	US-09-949-016-27127	Sequence 27127, A
17	77.8	601	4	US-09-949-016-35366	Sequence 35366, A
18	77.8	601	4	US-09-949-016-72490	Sequence 72490, A
19	77.8	601	4	US-09-949-016-102817	Sequence 102817, A
20	77.8	601	4	US-09-949-016-102836	Sequence 102836, A
21	77.8	601	4	US-09-949-016-160521	Sequence 160521, A
22	77.8	601	4	US-09-949-016-203142	Sequence 203142, A
23	77.8	601	4	US-09-949-016-203184	Sequence 203184, A
24	77.8	1001	4	US-09-671-317-418	Sequence 418, App
25	77.8	1797	4	US-09-891-641-34	Sequence 34, Appl
26	77.8	2465	1	US-08-421-661-5	Sequence 5, Appl
27	77.8	2634	4	US-09-949-016-4496	Sequence 4496, Ap

C	28	14	77.8	2808	4	US-09-917-254-27	Sequence 27, Appl
	29	14	77.8	4181	1	US-07-670-611-1	Sequence 1, Appl
	30	14	77.8	4181	1	US-08-220-674-1	Sequence 1, Appl
	31	14	77.8	4181	1	US-08-445-186-1	Sequence 1, Appl
	32	14	77.8	4181	1	US-08-446-549-1	Sequence 1, Appl
	33	14	77.8	4181	2	US-08-446-550-1	Sequence 1, Appl
	34	14	77.8	5385	4	US-09-920-804-1	Sequence 2, Appl
C	35	14	77.8	5661	3	US-08-938-105-2	Sequence 2, Appl
	36	14	77.8	11726	4	US-09-949-016-16968	Sequence 16968, A
	37	14	77.8	11726	4	US-09-949-016-16969	Sequence 16969, A
C	38	14	77.8	12725	4	US-09-949-016-14177	Sequence 14177, A
	39	14	77.8	16341	4	US-09-949-016-14579	Sequence 14579, A
	40	14	77.8	16341	4	US-09-949-016-14580	Sequence 14580, A
C	41	14	77.8	20063	4	US-09-949-016-14176	Sequence 14176, A
	42	14	77.8	20520	4	US-09-949-016-14394	Sequence 14394, A
	43	14	77.8	21784	4	US-09-820-002-3	Sequence 3, Appl
	44	14	77.8	29927	4	US-09-949-016-11814	Sequence 11814, A
	45	14	77.8	29927	4	US-09-949-016-17474	Sequence 17474, A

ALIGNMENTS

```
RESULT 1
US-09-872-051-11
; Sequence 11, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32 (nK603) and Composition and Methods for Det
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(18)
; OTHER INFORMATION: Zea maize genomic DNA and vector DNA
US-09-872-051-11

Query Match          100.0%; Score 18; DB 4; Length 18;
Best local similarity 100.0%; Pred. No. 0.54;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGCCTCTCTGCTGACTTT 18
Db 1 TGCCTCTCTGCTGACTTT 18

RESULT 2
US-09-872-051-8
; Sequence 8, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
```

```
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMG32 (nk603) and Composition and Methods for Detecting
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8
```

```
Query Match      100.0%; Score 18; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  TGCTGTTCTGCTGACTT 18
          |||||
Db      678  TGCTGTTCTGCTGACTT 635
```

```
RESULT 3
US-09-949-016-81715/c
; Sequence 81715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81715
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-81715
```

```
Query Match      88.9%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  TGCTGTTCTGCTGACT 16
          |||||
Db      349  TGCTGTTCTGCTGACT 334
```

```
RESULT 4
US-09-949-016-81717/c
```

```
; Sequence 81717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81717
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-81717
```

```
Query Match      88.9%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  TGCTGTTCTGCTGACT 16
          |||||
Db      223  TGCTGTTCTGCTGACT 208
```

```
RESULT 5
US-09-949-016-14319
; Sequence 14319, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14319
; LENGTH: 31391
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (1)..(31391)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14319
```

```
Query Match      88.9%; Score 16; DB 4; Length 31391;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  TGCTGTTCTGCTGACT 16
          |||||
Db      21183  TGCTGTTCTGCTGACT 21198
```

```
RESULT 6
US-09-949-016-14120/c
; Sequence 14120, Application US/09949016
```

```
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14120
LENGTH: 125192
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(125192)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14120
```

```
Query Match      88.3%; Score 16; DB 4; Length 125192;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGCTGTTCTGCTGACT 16
DB      48907 TGCTGTTCTGCTGACT 48892
```

```
RESULT 7
US-09-422-978-7605
Sequence 7605, Application US/09422978
Patent No. 6537751
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
FILE REFERENCE: GENSET.020CP1
CURRENT APPLICATION NUMBER: US/09/422,978
CURRENT FILING DATE: 1999-10-20
EARLIER APPLICATION NUMBER: US 09/298,850
EARLIER FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,732
EARLIER FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: US 60/082,614
EARLIER FILING DATE: 1998-04-21
NUMBER OF SEQ ID NOS: 11796
SEQ ID NO 7605
LENGTH: 21
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1..21
OTHER INFORMATION: upstream amplification primer 99-9623 for SEQ 3671,
US-09-422-978-7605
```

```
Query Match      83.3%; Score 15; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CTGTTCTGCTGACTT 17
DB      2 CTGTTCTGCTGACTT 16
```

```
RESULT 8
US-09-949-016-58184
Sequence 58184, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58184
LENGTH: 601
TYPE: DNA
ORGANISM: Human
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-58184
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 TGTTCTGCTGACTTT 18
DB      512 TGTTCTGCTGACTTT 526
```

```
RESULT 9
US-09-949-016-15108
Sequence 15108, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15108
LENGTH: 21125
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(21125)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15108
```

```
Query Match      83.3%; Score 15; DB 4; Length 21125;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CTGTTCTGCTGACTT 17
DB      11408 CTGTTCTGCTGACTT 11422
```

RESULT 10

US-09-949-016-13429/c
; Sequence 13429, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13429
; LENGTH: 57761
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(57761)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13429

Query Match 83.3%; Score 15; DB 4; Length 57761;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 TGTTCTGCTGACTTT 18
Db 27356 TGTTCTGCTGACTTT 27342

RESULT 11
US-09-198-452A-5457/c
; Sequence 5457, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 5457
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-5457

Query Match 77.8%; Score 14; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 GTTCTGCTGACTTT 18
Db 16 GTTCTGCTGACTTT 3

RESULT 12
US-09-016-434-403/c
; Sequence 403, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 403:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT07
; CLONE: 2121175
US-09-016-434-403

Query Match 77.8%; Score 14; DB 4; Length 241;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGCTGTTCTGCTGA 14
Db 163 TGCTGTTCTGCTGA 150

RESULT 13
US-09-621-976-9793/c
; Sequence 9793, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Giordano, J.Y.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9793
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9793

Query Match 77.8%; Score 14; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TGCTCTGCTGACTT 17
|||||
DB 85 TGCTCTGCTGACTT 72

RESULT 14
US-09-489-039A-4
; Sequence 4, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4

Query Match 77.8%; Score 14; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGA 14
|||||
DB 258 TGCTGTTCTGCTGA 271

RESULT 15
US-09-949-016-19121/C
; Sequence 19121, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19121
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19121

Query Match 77.8%; Score 14; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGA 14
|||||
DB 553 TGCTGTTCTGCTGA 540

Search completed: February 10, 2005, 08:50:03
Job time : 7.5316 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 10, 2005, 07:18:58 ; Search time 5.5316 Seconds
(Without alignments)
5324.730 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 18

Sequence: 1 taccacgcgacacacttc 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patente NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	4 US-09-872-051-10	Sequence 10, Appl
2	18	100.0	1183	4 US-09-872-051-8	Sequence 8, Appl
3	14	77.8	174170	4 US-09-949-016-14810	Sequence 14810, A
4	14	77.8	174170	4 US-09-949-016-14811	Sequence 14811, A
5	14	77.8	174318	4 US-09-949-016-11880	Sequence 11880, A
6	14	77.8	174318	4 US-09-949-016-14812	Sequence 14812, A
7	14	77.8	174318	4 US-09-949-016-14813	Sequence 14813, A
8	13	72.2	296	4 US-09-313-294A-4452	Sequence 4452, Ap
9	13	72.2	321	4 US-09-270-767-29026	Sequence 29026, A
10	13	72.2	1419	4 US-09-270-767-13127	Sequence 13127, A
11	13	72.2	35081	2 US-08-752-760A-1	Sequence 1, Appl
12	13	72.2	4403765	3 US-09-103-840A-2	Sequence 2, Appl
13	13	72.2	4411529	3 US-09-103-840A-1	Sequence 1, Appl
14	12	66.7	84	4 US-09-513-999C-15367	Sequence 15367, A
15	12	66.7	262	4 US-09-513-999C-14094	Sequence 14094, A
16	12	66.7	435	4 US-09-270-767-6685	Sequence 6685, Ap
17	12	66.7	435	4 US-09-270-767-21967	Sequence 21967, A
18	12	66.7	477	4 US-09-270-767-4902	Sequence 4902, Ap
19	12	66.7	477	4 US-09-270-767-20184	Sequence 20184, A
20	12	66.7	548	4 US-09-791-105B-31	Sequence 31, Appl
21	12	66.7	558	4 US-09-902-540-1969	Sequence 1969, Ap
22	12	66.7	601	4 US-09-949-016-57353	Sequence 57353, A
23	12	66.7	601	4 US-09-949-016-205395	Sequence 205395, A
24	12	66.7	691	3 US-08-998-416-742	Sequence 742, App
25	12	66.7	712	3 US-08-998-416-827	Sequence 827, Appl
26	12	66.7	813	3 US-08-818-112-3	Sequence 3, Appl
27	12	66.7	813	3 US-08-818-111-3	Sequence 3, Appl

28	12	66.7	813	3 US-09-056-556-3	Sequence 3, Appl
29	12	66.7	813	3 US-09-072-596-3	Sequence 3, Appl
30	12	66.7	813	4 US-09-072-967-3	Sequence 3, Appl
31	12	66.7	1378	4 US-09-621-377B-1	Sequence 1, Appl
32	12	66.7	1378	4 US-10-031-893-1	Sequence 5, Appl
33	12	66.7	1391	2 US-08-813-940-5	Sequence 1194, A
34	12	66.7	1414	4 US-09-270-767-11194	Sequence 11259, A
35	12	66.7	1497	4 US-09-270-767-11259	Sequence 13, Appl
36	12	66.7	1539	4 US-09-716-865-13	Sequence 1, Appl
37	12	66.7	1668	2 US-08-901-547A-1	Sequence 1, Appl
38	12	66.7	1859	3 US-07-861-458C-1	Sequence 3, Appl
39	12	66.7	2558	3 US-08-999-733-3	Sequence 108, App
40	12	66.7	3093	4 US-09-614-221A-108	Sequence 230, App
41	12	66.7	3483	4 US-09-711-164-230	Sequence 49, Appl
42	12	66.7	4175	1 US-08-106-691B-49	Sequence 1, Appl
43	12	66.7	4175	3 US-08-202-841A-1	Sequence 84, Appl
44	12	66.7	4175	5 PCT-US93-06251-84	Sequence 14, Appl
45	12	66.7	9408	4 US-09-418-710-14	

ALIGNMENTS

```
RESULT 1
US-09-872-051-10
; Sequence 10, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMG32 (ink603) and Composition and Methods for Det
; FILE REFERENCE: 38-21 (52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(18)
; OTHER INFORMATION: zea maize plastid DNA and vector DNA
US-09-872-051-10
Query Match 100.0%; Score 18; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 TACCACGCGACACCTTC 18
Db 1 TACCACGCGACACCTTC 18
RESULT 2
US-09-872-051-8
; Sequence 8, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
```

```
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMG32 (nk603) and Composition and Methods for Detecting
TITLE OF INVENTION: Thereof
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1183
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1183)
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
OTHER INFORMATION: 165-381 construct vector DNA
OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8
```

```
Query Match      100.0%; Score 18; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 TACCGCGACACTTC 18
      |||||
Db      373 TACCGCGACACTTC 390
```

```
RESULT 3
US-09-949-016-14810
Sequence 14810, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14810
LENGTH: 174170
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(174170)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14810
```

```
Query Match      77.8%; Score 14; DB 4; Length 174170;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      4 CACCGGACACTT 17
      |||||
Db      86907 CACCGGACACTT 86920
```

```
RESULT 4
US-09-949-016-14811
Sequence 14811, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14811
LENGTH: 174170
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(174170)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14811
```

```
Query Match      77.8%; Score 14; DB 4; Length 174170;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      4 CACCGGACACTT 17
      |||||
Db      86907 CACCGGACACTT 86920
```

```
RESULT 5
US-09-949-016-11880
Sequence 11880, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11880
LENGTH: 174318
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(174318)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11880
```

```
Query Match      77.8%; Score 14; DB 4; Length 174318;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 4 CACGGCAGACACTT 17
| | | | | | | | | |
Db 87055 CACGGCAGACACTT 87068

RESULT 6
US-09-949-016-14812

; Sequence 14812, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14812
; LENGTH: 174318
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174318)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14812

Query Match 77.8%; Score 14; DB 4; Length 174318;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGGCAGACACTT 17
| | | | | | | | | |
Db 87055 CACGGCAGACACTT 87068

RESULT 7
US-09-949-016-14813
; Sequence 14813, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/221,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14813
; LENGTH: 174318
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174318)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14813

Query Match 77.8%; Score 14; DB 4; Length 174318;

Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGGCAGACACTT 17
| | | | | | | | | |
Db 87055 CACGGCAGACACTT 87068

RESULT 8

US-09-313-294A-4452
; Sequence 4452, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4452
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348447H1
US-09-313-294A-4452

Query Match 72.2%; Score 13; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACCACCGCAGACA 14
| | | | | | | | | |
Db 260 ACCACCGCAGACA 272

RESULT 9
US-09-270-767-29026/c
; Sequence 29026, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29026
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29026

Query Match 72.2%; Score 13; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGGCAGACAC 15
| | | | | | | | | |
Db 62 CCACGGCAGACAC 50

RESULT 10
US-09-270-767-13127/c
; Sequence 13127, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13127
LENGTH: 1419
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-13127

Query Match 72.2%; Score 13; DB 4; Length 1419;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCGCGCACACTT 15
DB 536 CCGCGCACACTT 524

RESULT 11
US-08-752-760A-1
Sequence 1, Application US/08752760A
Patent No. 5877011
GENERAL INFORMATION:
APPLICANT: Armentano, Donna
APPLICANT: Gregory, Richard J.
APPLICANT: Smith, Alan E.
TITLE OF INVENTION: CHIMERIC ADENOIRAL VECTORS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Baker & Botts, L.L.P.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,760A
FILING DATE: 20-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Seide, Rochelle K
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: A31385
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-705-5000
TELEFAX: 212-705-5020
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 35081 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-752-760A-1

Query Match 72.2%; Score 13; DB 2; Length 35081;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCGCGCACACTT 18
DB 6154 CCGCGCACACTT 6166

RESULT 12
US-09-103-840A-2/C
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Query Match 72.2%; Score 13; DB 3; Length 4403765;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ACGCGACACTT 17
DB 1224061 ACGCGACACTT 1224049

RESULT 13
US-09-103-840A-1/C
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 72.2%; Score 13; DB 3; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ACGCGACACTT 17
DB 1224533 ACGCGACACTT 1224521

RESULT 14
US-09-513-999C-15367/C
Sequence 15367, Application US/09513999C
Patent No. 6783961

```

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15367
; LENGTH: 84
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15367

```

```

Query Match      66.7%; Score 12; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      7 GCGACACACTTC 18
         |||||
Db      16 GCGACACACTTC 5

```

```

RESULT 15
US-09-513-999C-14094/C
; Sequence 14094, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14094
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 35
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 39
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 41
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 50
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 262
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-14094

```

```

Query Match      66.7%; Score 12; DB 4; Length 262;
Best Local Similarity 100.0%; Pred. No. 2e+02;

```

```

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2 ACCACGCGACAC 13
         |||||
Db      199 ACCACGCGACAC 188

```

```

Search completed: February 10, 2005, 08:50:01
Job time : 18.5314 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using BW model

Run on: February 10, 2005, 07:47:34 ; Search time 111.027 Seconds
(without alignments)
7855.666 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 18

Sequence: 1 taccacgcgacacacttc 18

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:*
1: gb_Da:*
2: gb_Htg:*
3: gb_In:*
4: gb_Om:*
5: gb_Ov:*
6: gb_Pat:*
7: gb_Pn:*
8: gb_Pl:*
9: gb_Pr:*
10: gb_Ro:*
11: gb_Sts:*
12: gb_Sy:*
13: gb_Un:*
14: gb_Vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	AX342371	Sequence
2	18	100.0	18	AX342369	Sequence
3	17	94.4	865	AX325002	Sequence
4	16	88.9	98319	AX090523	Sequence
5	15	83.3	363	BV074353	Sequence
6	15	83.3	747	AX065627	Sequence
7	15	83.3	69794	AC013773	Sequence
8	15	83.3	89474	AC013802	Sequence
9	15	83.3	110000	AP006502_01	Sequence
10	15	83.3	170777	AC091307	Sequence
11	15	83.3	299587	AC091307	Sequence
12	15	83.3	309400	AX127153	Sequence
13	15	83.3	325651	AP005283	Sequence
14	15	83.3	349115	AX127156	Sequence
15	15	83.3	349980	AX127152	Sequence
16	14	77.8	435	CO499894	Sequence
17	14	77.8	435	CO508863	Sequence
18	14	77.8	435	AX308198	Sequence
19	14	77.8	437	CO469529	Sequence

20	14	77.8	484	6	CO478697	Sequence
21	14	77.8	1889	9	AF383169	Sequence
22	14	77.8	2000	6	AX655473	Sequence
23	14	77.8	3154	5	BC044418	Sequence
24	14	77.8	5813	6	CO580220	Sequence
25	14	77.8	1111	1	AE007286	Sequence
26	14	77.8	29989	3	U40059	Sequence
27	14	77.8	71018	2	AC131498	Sequence
28	14	77.8	90027	2	AC079522	Sequence
29	14	77.8	96438	2	AC150979	Sequence
30	14	77.8	97538	2	AC017943	Sequence
31	14	77.8	105805	2	AC022280	Sequence
32	14	77.8	110000	1	AE017283_09	Sequence
33	14	77.8	110000	2	LMFCHR36_10	Sequence
34	14	77.8	112673	2	AY648026	Sequence
35	14	77.8	118020	2	AC134824	Sequence
36	14	77.8	129402	2	AP003341	Sequence
37	14	77.8	131054	8	AP003341	Sequence
38	14	77.8	135611	14	AF169823	Sequence
39	14	77.8	144023	9	AC083928	Sequence
40	14	77.8	156889	2	AC004586	Sequence
41	14	77.8	158084	8	AP003492	Sequence
42	14	77.8	158594	2	AC087768	Sequence
43	14	77.8	165083	10	AC120554	Sequence
44	14	77.8	165497	5	BX511107	Sequence
45	14	77.8	170939	3	AC007757	Sequence

ALIGNMENTS

RESULT 1	AX342371	18 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342371	Sequence 10 from Patent EP1167531.			
DEFINITION	AX342371	AX342371.1 GI:18151814			
ACCESSION	AX342371				
VERSION	AX342371.1	GI:18151814			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Behr, C.F., Hironaka, C., Heck, G.R. and You, J.				
TITLE	Corn transformant pV-zmgf32 (nke03) and compositions and methods for detection thereof				
JOURNAL	Patent: EP 1167531-A 10 02-JAN-2002;				
FEATURES	Monsanto Technology LLC (US)				
SOURCE	Location/Qualifiers				
	1..18				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="zea maize plastid DNA and vector DNA"				
ORIGIN					
Query Match	100.0%;	Score 18;	DB 6;	Length 18;	
Best Local Similarity	100.0%;	Pred. No. 0.57;			
Matches	18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1 TACCACGCGACACACTTC 18				
Db	1 TACCACGCGACACACTTC 18				
RESULT 2	AX342369	1183 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342369	Sequence 8 from Patent EP1167531.			
DEFINITION	AX342369				
ACCESSION	AX342369				
VERSION	AX342369.1	GI:18151812			
KEYWORDS					
SOURCE					
ORGANISM					
	synthetic construct				
	synthetic construct				

other sequences; artificial sequences.

REFERENCE 1
 Behr, C.F., Hironaka, C., Heck, G.R. and You, J.
 TITLE Corn transformant pv-zmgc32 (mk603) and compositions and methods for detection thereof
 JOURNAL Patent: EP 1167531-A 8 02-JAN-2002;
 Monsanto Technology LLC (US)
 FEATURES
 source location/Qualifiers
 1. 1183
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="1-164 Agrobacterium tumefaciens nos 3' terminator 165-381 construct vector DNA 382-686 Zea maize plasmid genes, rps11 and rps4 687-1183 Zea maize genomic DNA"

ORIGIN
 Query Match 100.0%; Score 18; DB 6; Length 1183;
 Best Local Similarity 100.0%; Pred. No. 0.38;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGGACACACTTC 18
 |||||
 373 TACCACGGACACACTTC 390

RESULT 3
 HSA325002 865 bp DNA linear PRI 18-JUL-2002
 LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone
 DEFINITION NL1-DK11R.
 ACCESSION AJ325002
 VERSION AJ325002.1 GI:15869396
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 865)
 Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Khasna, S.M., Podowski, R.M., Matshkin, Y.G., Gyanchandani, A., Muravenko, O.V., Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I., Kisselev, L.L., Wasserman, W., Wahlestedt, C. and Zabarovsky, E.R.
 TITLE Not1 flanking sequences: a tool for gene discovery and verification of the human genome
 JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
 MEDLINE 22131767
 PUBMED 12136098
 2 (bases 1 to 865)
 Zabarovsky, E.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden

FEATURES
 source location/Qualifiers
 1. 865
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="NL1-DK11R"

ORIGIN
 Query Match 94.4%; Score 17; DB 9; Length 865;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACCACGGACACACTTC 18
 |||||
 546 ACCACGGACACACTTC 562

RESULT 4
 AC090523

LOCUS AC090523 98319 bp DNA linear INV 27-FEB-2001
 DEFINITION Caenorhabditis briggsae cosmid CB015N01, complete sequence.
 ACCESSION AC090523
 VERSION AC090523.1 GI:13129526
 KEYWORDS HTG.
 SOURCE Caenorhabditis briggsae
 ORGANISM Caenorhabditis briggsae
 Eukaryota; Metazoa; Nematoea; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 98319)
 Washington University Genome Sequencing Center.
 TITLE The C. briggsae Genome Sequencing Project
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 98319)
 REFERENCE Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-2001) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63110, USA
 e-mail: jepiet@wustl.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES
 source location/Qualifiers
 1. 98319
 /organism="Caenorhabditis briggsae"
 /mol_type="genomic DNA"
 /strain="Gujarat G16"
 /db_xref="taxon:6238"
 /clone="CB015N01"
 /product="cDNA-Len"
 /note="codon recognized: CUG"
 /product="cDNA-Pro"
 /note="codon recognized: CCA"

ORIGIN
 Query Match 88.9%; Score 16; DB 3; Length 98319;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGGACACACTTC 18
 |||||
 20762 CCACGGACACACTTC 20777

RESULT 5
 BV074353 363 bp DNA linear STS 31-MAY-2003
 LOCUS S212P6009FB1.T0 CZECHII/El Mus musculus STS genomic, sequence
 DEFINITION tagged site.
 ACCESSION BV074353
 VERSION BV074353.1 GI:31190148
 KEYWORDS STS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 363)
 Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
 TITLE The mosaic structure of variation in the laboratory mouse genome
 JOURNAL Nature 420 (6915), 574-578 (2002)
 MEDLINE 22354684
 PUBMED 12466852
 COMMENT

Contact: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome
 Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kersti@genome.wi.mit.edu

Primer A: None
 Primer B: None
 STS size: 363

Protocol:
 WGS-discovery: Paired-end low-coverage whole genome shotgun reads
 were generated from 129S1/SvImt, C3H/HeJ, and BALB/cByJ. The WGS
 reads were placed uniquely on the MGSv3 C57BL/6J assembly and SNP
 detection was carried out by SSMHA-SNP. 225,000 reads were
 annotated
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
 and the strain from which the particular read came. The validation
 rate for these SNPs was estimated at approximately 98%.

FEATURES

source

1. 363
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /map="16 27-306 28672947-28672668"
 /clone_lib="CZECCHI/El"
 <1..>363

ORIGIN

STS

Query Match 83.3%; Score 15; DB 11; Length 363;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

4 CACGGCAGACACTTC 18
 |||||
 190 CACGGCAGACACTTC 204

RESULT 6

AX065627/c 747 bp DNA linear PAT 24-JAN-2001
 LOCUS Sequence 753 from Patent WO0100844.
 AX065627
 AX065627.1 GI:12543339

SOURCE

Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriales; Corynebacterium.

REFERENCE

1 Pompejus, M., Kroege, B., Schroeder, H., Zeidler, O. and Habermann, G.
 corynebacterium glutamicum genes encoding proteins involved in
 carbon metabolism and energy production
 Patent: WO 0100844-A 753 04-JAN-2001;
 BASF AKTIEGENSCHAFT (DE)

FEATURES

source

CDS

1. 747
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..727
 /note="unamed protein product; RXN02554"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC26053.1"
 /db_xref="GI:12543340"
 /translation="MSHTKPSIALIGAGRVSSLSARSAVAAGYEVKAGSGAVDKIAL
 TAILMPCAVSTADQAKADIVFLAHLKFRSVNPATLGGKIVIDTMHWVPVNG
 ELERIDDPSTSEILAEFFAGSTWVSFNIHGYEIRDAGTGALAIAYTDVDVAGA
 OVAOLISFGVPLNIGALENGRIIEPQGEAFGAILNDSLELVNR"

ORIGIN

Query Match 83.3%; Score 15; DB 6; Length 747;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGGCAGACACTTC 18
 |||||
 Db 21 CACGGCAGACACTTC 7

RESULT 7

AC013773/c 69794 bp DNA linear HTG 13-JUL-2000
 LOCUS Homo sapiens clone RP11-10A24, LOW-PASS SEQUENCE SAMPLING.
 AC013773
 AC013773.2 GI:9123806
 HTG; HTGS_PHASE0.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 TITLE Unpublished
 JOURNAL
 REFERENCE
 AUTHORS

1 (bases 1 to 69794)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 2 (bases 1 to 69794)
 Homo sapiens, clone RP11-10A24
 Unpublished
 2 (bases 1 to 69794)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouhgalter, B.,
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., Dearielano, K., Dewar, K., Domito, M., Donelan, L., Doyle, M.,
 Ferreira, P., Fitzhugh, M., Forrest, C., Funke, R., Gage, D.,
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
 Lehoczy, J., Liu, C., Locke, K., MacDonald, P., Marquis, N.,
 McEwan, P., McGuirk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6425747.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WtBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L3033
 Center clone name: 10_A_24

NOTE: This record contains 77 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 792: contig of 792 bp in length
 * 793 792: gap of 100 bp
 * 893 1698: contig of 806 bp in length
 * 1699 1798: gap of 100 bp
 * 1799 2602: contig of 804 bp in length
 * 2603 2702: gap of 100 bp
 * 2703 3518: contig of 816 bp in length

* 3519 3618: gap of 100 bp
* 3619 4415: contig of 797 bp in length
* 4416 4515: gap of 100 bp
* 4516 5386: contig of 870 bp in length
* 5386 5485: gap of 100 bp
* 5485 6305: contig of 820 bp in length
* 6305 6405: gap of 100 bp
* 6405 7233: contig of 828 bp in length
* 7233 7334: gap of 100 bp
* 7334 8139: contig of 806 bp in length
* 8139 8239: gap of 100 bp
* 8239 9026: contig of 787 bp in length
* 9026 9126: gap of 100 bp
* 9126 9933: contig of 806 bp in length
* 9933 10032: gap of 100 bp
* 10032 10840: contig of 808 bp in length
* 10840 10940: gap of 100 bp
* 10940 11753: contig of 813 bp in length
* 11753 11853: gap of 100 bp
* 11853 12657: contig of 804 bp in length
* 12657 12757: gap of 100 bp
* 12757 13574: contig of 817 bp in length
* 13574 13674: gap of 100 bp
* 13674 14456: contig of 782 bp in length
* 14456 14556: gap of 100 bp
* 14556 15353: contig of 797 bp in length
* 15353 15453: gap of 100 bp
* 15453 16298: contig of 845 bp in length
* 16298 16398: gap of 100 bp
* 16398 17218: contig of 820 bp in length
* 17218 17318: gap of 100 bp
* 17318 18115: contig of 798 bp in length
* 18115 18215: gap of 100 bp
* 18215 19012: contig of 796 bp in length
* 19012 19112: gap of 100 bp
* 19112 19912: contig of 800 bp in length
* 19912 20012: gap of 100 bp
* 20012 20835: contig of 823 bp in length
* 20835 20935: gap of 100 bp
* 20935 21731: contig of 796 bp in length
* 21731 21831: gap of 100 bp
* 21831 22603: contig of 772 bp in length
* 22603 22703: gap of 100 bp
* 22703 23507: contig of 804 bp in length
* 23507 23607: gap of 100 bp
* 23607 24420: contig of 813 bp in length
* 24420 24520: gap of 100 bp
* 24520 25341: contig of 821 bp in length
* 25341 25441: gap of 100 bp
* 25441 26235: contig of 794 bp in length
* 26235 26335: gap of 100 bp
* 26335 27114: contig of 779 bp in length
* 27114 27214: gap of 100 bp
* 27214 27997: contig of 783 bp in length
* 27997 28097: gap of 100 bp
* 28097 28923: contig of 826 bp in length
* 28923 29023: gap of 100 bp
* 29023 29844: contig of 821 bp in length
* 29844 29944: gap of 100 bp
* 29944 30746: contig of 802 bp in length
* 30746 30846: gap of 100 bp
* 30846 31663: contig of 817 bp in length
* 31663 31763: gap of 100 bp
* 31763 32575: contig of 812 bp in length
* 32575 32675: gap of 100 bp
* 32675 33491: contig of 816 bp in length
* 33491 33591: gap of 100 bp
* 33591 34407: contig of 816 bp in length
* 34407 34507: gap of 100 bp
* 34507 35309: contig of 802 bp in length
* 35309 35409: gap of 100 bp
* 35409 36196: contig of 787 bp in length
* 36196 36296: gap of 100 bp

* 36297 37109: contig of 813 bp in length
* 37109 37209: gap of 100 bp
* 37209 38023: contig of 813 bp in length
* 38023 38122: gap of 100 bp
* 38122 38917: contig of 795 bp in length
* 38917 39018: gap of 100 bp
* 39018 39838: contig of 821 bp in length
* 39838 39939: gap of 100 bp
* 39939 40723: contig of 784 bp in length
* 40723 40822: gap of 100 bp
* 40822 41610: contig of 788 bp in length
* 41610 41710: gap of 100 bp
* 41710 42548: contig of 838 bp in length
* 42548 42649: gap of 100 bp
* 42649 43469: contig of 821 bp in length
* 43469 43569: gap of 100 bp
* 43569 44387: contig of 818 bp in length
* 44387 44487: gap of 100 bp
* 44487 45308: contig of 821 bp in length
* 45308 45408: gap of 100 bp
* 45408 46219: contig of 811 bp in length
* 46219 46319: gap of 100 bp
* 46319 47123: contig of 804 bp in length
* 47123 47223: gap of 100 bp
* 47223 48019: contig of 796 bp in length
* 48019 48119: gap of 100 bp
* 48119 48955: contig of 837 bp in length
* 48955 49056: gap of 100 bp
* 49056 49826: contig of 770 bp in length
* 49826 49926: gap of 100 bp
* 49926 50740: contig of 814 bp in length
* 50740 50840: gap of 100 bp
* 50840 51638: contig of 798 bp in length
* 51638 51738: gap of 100 bp
* 51738 52551: contig of 813 bp in length
* 52551 52652: gap of 100 bp
* 52652 53463: contig of 812 bp in length
* 53463 53563: gap of 100 bp
* 53563 54355: contig of 793 bp in length
* 54355 54456: gap of 100 bp
* 54456 55285: contig of 829 bp in length
* 55285 55385: gap of 100 bp
* 55385 56164: contig of 779 bp in length
* 56164 56264: gap of 100 bp
* 56264 57080: contig of 816 bp in length
* 57080 57180: gap of 100 bp
* 57180 57964: contig of 784 bp in length
* 57964 58064: gap of 100 bp
* 58064 58916: contig of 852 bp in length
* 58916 59016: gap of 100 bp
* 59016 59868: contig of 852 bp in length
* 59868 59968: gap of 100 bp
* 59968 60728: contig of 760 bp in length
* 60728 60829: gap of 100 bp
* 60829 61630: contig of 802 bp in length
* 61630 61730: gap of 100 bp
* 61730 62566: contig of 836 bp in length
* 62566 62667: gap of 100 bp
* 62667 63483: contig of 817 bp in length
* 63483 63583: gap of 100 bp
* 63583 64376: contig of 793 bp in length
* 64376 64477: gap of 100 bp
* 64477 65284: contig of 808 bp in length
* 65284 65384: gap of 100 bp
* 65384 66185: contig of 801 bp in length

Query Match 83.3%; Score 15; DB 2; Length 69794;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 24772 CCACGCGACACTT 24758
CCACGCGACACTT 17
|||||

RESULT 8
AC013802
LOCUS
DEFINITION
Homo sapiens clone RP11-20M13, LOW-PASS SEQUENCE SAMPLING.
AC013802
AC013802.3 GI:9123968
VERSION
HTG: HTGS PHASE0.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 89474)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-20M13
Unpublished
2 (bases 1 to 89474)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barina, N., Beckerly, R., Boguslavsky, L., Bouknight, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Lehoczky, J., Liu, C., Locks, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tittell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6957824.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3932
Center clone name: 20_M_13

* NOTE: This record contains 103 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
726: contig of 726 bp in length
727 826: gap of 100 bp
728 827: gap of 100 bp
1515: contig of 689 bp in length
1516 1615: gap of 100 bp
1616 2345: contig of 730 bp in length
2346 2445: gap of 100 bp
2446 3154: contig of 709 bp in length
3155 3254: gap of 100 bp
3255 3983: contig of 729 bp in length
3984 4083: gap of 100 bp
4084 4810: contig of 727 bp in length
4811 4911: gap of 100 bp
4911 5625: contig of 715 bp in length
5626 5725: gap of 100 bp

5726 6448: contig of 723 bp in length
6449 6548: gap of 100 bp
6549 7262: contig of 714 bp in length
7263 7362: gap of 100 bp
7363 8036: contig of 674 bp in length
8037 8136: gap of 100 bp
8137 8854: contig of 718 bp in length
8855 9694: gap of 100 bp
9695 9794: contig of 740 bp in length
9795 9994: gap of 100 bp
10508 10507: contig of 713 bp in length
10508 10607: gap of 100 bp
10608 11333: contig of 726 bp in length
11334 11433: gap of 100 bp
11434 12142: contig of 709 bp in length
12143 12242: gap of 100 bp
12243 12998: contig of 756 bp in length
12999 13098: gap of 100 bp
13099 13826: contig of 728 bp in length
13827 13926: gap of 100 bp
13927 14647: contig of 721 bp in length
14648 14747: gap of 100 bp
14748 15449: contig of 702 bp in length
15450 15549: gap of 100 bp
15550 16249: contig of 700 bp in length
16250 16349: gap of 100 bp
16350 17058: contig of 709 bp in length
17059 17158: gap of 100 bp
17159 17864: contig of 706 bp in length
17865 17964: gap of 100 bp
17965 18697: contig of 733 bp in length
18698 18797: gap of 100 bp
18798 19512: contig of 715 bp in length
19513 19612: gap of 100 bp
19613 20338: contig of 726 bp in length
20339 20438: gap of 100 bp
20439 21177: contig of 739 bp in length
21178 21277: gap of 100 bp
21278 22000: contig of 723 bp in length
22001 22100: gap of 100 bp
22101 22816: contig of 716 bp in length
22817 22916: gap of 100 bp
22917 23627: contig of 711 bp in length
23628 23727: gap of 100 bp
23728 24461: contig of 734 bp in length
24462 24561: gap of 100 bp
24562 25276: contig of 715 bp in length
25277 25376: gap of 100 bp
25377 26097: contig of 721 bp in length
26098 26197: gap of 100 bp
26198 26908: contig of 711 bp in length
26909 27008: gap of 100 bp
27009 27738: contig of 730 bp in length
27739 27838: gap of 100 bp
27839 28561: contig of 723 bp in length
28562 28661: gap of 100 bp
28662 29375: contig of 714 bp in length
29376 29475: gap of 100 bp
29476 30189: contig of 714 bp in length
30189 30289: gap of 100 bp
30290 30997: contig of 708 bp in length
30998 31097: gap of 100 bp
31098 31820: contig of 723 bp in length
31821 31920: gap of 100 bp
31921 32641: contig of 721 bp in length
32642 32741: gap of 100 bp
32742 33469: contig of 728 bp in length
33470 33569: gap of 100 bp
33570 34286: contig of 717 bp in length
34287 34386: gap of 100 bp
34387 35091: contig of 705 bp in length
35092 35191: gap of 100 bp
35192 35901: contig of 710 bp in length

```
* 35902 36001: gap of 100 bp
* 36002 36718: contig of 717 bp in length
* 36719 36818: gap of 100 bp
* 36819 37540: contig of 722 bp in length
* 37541 37640: gap of 100 bp
* 37641 38361: contig of 721 bp in length
* 38362 38461: gap of 100 bp
* 38462 39175: contig of 714 bp in length
* 39176 39275: gap of 100 bp
* 39276 39981: contig of 706 bp in length
* 39982 40081: gap of 100 bp
* 40082 40802: contig of 721 bp in length
* 40803 40902: gap of 100 bp
* 40903 41593: contig of 691 bp in length
* 41594 42434: gap of 100 bp
* 42435 42534: gap of 100 bp
* 42535 43282: contig of 748 bp in length
* 43283 43383: gap of 100 bp
* 43384 44064: contig of 682 bp in length
* 44065 44861: contig of 697 bp in length
* 44862 44961: gap of 100 bp
* 44962 45684: contig of 723 bp in length
* 45685 45784: gap of 100 bp
* 45785 46503: contig of 719 bp in length
* 46504 46603: gap of 100 bp
* 46604 47311: contig of 708 bp in length
* 47312 47411: gap of 100 bp
* 47412 48109: contig of 698 bp in length
* 48110 48209: gap of 100 bp
* 48210 48913: contig of 704 bp in length
* 48914 49013: gap of 100 bp
* 49014 49730: contig of 717 bp in length
* 49731 49830: gap of 100 bp
* 49831 50545: contig of 715 bp in length
* 50546 50645: gap of 100 bp
* 50646 51356: contig of 711 bp in length
* 51357 51456: gap of 100 bp
* 51457 52191: contig of 735 bp in length
* 52192 52291: gap of 100 bp
* 52292 53025: contig of 734 bp in length
* 53026 53125: gap of 100 bp
* 53126 53872: contig of 747 bp in length
* 53873 53972: gap of 100 bp
* 53973 54689: contig of 717 bp in length
* 54690 54789: gap of 100 bp
* 54790 55515: contig of 726 bp in length
* 55516 55615: gap of 100 bp
* 55616 56327: contig of 712 bp in length
* 56328 56427: gap of 100 bp
* 56428 57132: contig of 705 bp in length
* 57133 57232: gap of 100 bp
* 57233 57971: contig of 739 bp in length
* 57972 58071: gap of 100 bp
* 58072 58792: contig of 720 bp in length
* 58793 58891: gap of 100 bp
* 58892 59592: contig of 701 bp in length
```

Query Match 83.3%; Score 15; DB 2; Length 89474;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCGACACACTT 17
DB 49056 CCACGCGACACACTT 49070

RESULT 9
AP006502.01
WPCOMMENT
Sequence split into 17 fragments LOCUS AP006502 Accession AP006502
Fragment Name Begin End

```
AP006502.00 1 110000  
AP006502.01 100001 210000  
AP006502.02 200001 310000  
AP006502.03 300001 410000  
AP006502.04 400001 510000  
AP006502.05 500001 610000  
AP006502.06 600001 710000  
AP006502.07 700001 810000  
AP006502.08 800001 910000  
AP006502.09 900001 1010000  
AP006502.10 1000001 1110000  
AP006502.11 1100001 1210000  
AP006502.12 1200001 1310000  
AP006502.13 1300001 1410000  
AP006502.14 1400001 1510000  
AP006502.15 1500001 1610000  
AP006502.16 1600001 1628011  
Continuation (2 of 17) of AP006502 from base 100001 (AP006502 Cyanidioschyzon merolae 81
```

Query Match 83.3%; Score 15; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACAC 15
DB 83124 TACCACGCGACACAC 83138

RESULT 10

LOCUS AC148444
DEFINITION Rhinolophus ferrumequinum clone vM8C7-327616, WORKING DRAFT
SEQUENCE, 7 ordered pieces.
AC148444

ACCESSION AC148444.2 GI:45917393
VERSION
KEYWORDS HTGS, PHASE2, HTGS, DRAFT.

SOURCE Rhinolophus ferrumequinum (Greater horseshoe bat)
ORGANISM Rhinolophus ferrumequinum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chiroptera; Microchiroptera; Rhinolophidae;
Rhinolophinae; Rhinolophus.
1 (bases 1 to 170777)
Antoneilla, A., Ayle, K., Benjamin, B., Blakesley, R.W.,
Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
Coleman, H., Daki, N., Engle, J., Grant, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Huxle, B.,
Idol, J.R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P.,
Larson, S., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.U., Maduro, V.B.,
Margules, E.H., Mastello, C., Maekel, B., McDowell, J.,
Mullikin, J.C., Paguirigan, C., Portnoy, M.E., Prasad, A., Puri, O.,
Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C.,
Stratipoulos, S., Thomas, J.W., Thomas, P.J., Tsipouris, V., Vogt, J.L.,
Wetherby, K.D., Young, A. and Green, E.D.

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (04-MAR-2004) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 170777)
Green, E.D.
Direct Submission
Submitted (01-APR-2004) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
On Apr 1, 2004 this sequence version replaced gi:44917669.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@ngri.nih.gov
----- Project Information
Center project name: gcm

Center clone name: 327G16

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 169390 bases at least Q40
Consensus quality: 169842 bases at least Q30
Consensus quality: 170073 bases at least Q20
Insert size: 13300; agarose-fp
Insert size: 170177; sum-of-contigs
Quality coverage: 11.58x in Q20 bases; agarose-fp
Quality coverage: 9.05x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 17243: contig of 17243 bp in length
1 17244 17343: gap of unknown length
* 17344 24932: contig of 7589 bp in length
* 24933 25032: gap of unknown length
* 25033 75704: contig of 50672 bp in length
* 75705 75804: gap of unknown length
* 75805 95337: contig of 19533 bp in length
* 95338 95437: gap of unknown length
* 95438 101416: contig of 5779 bp in length
* 101417 101516: gap of unknown length
* 101517 163768: contig of 62252 bp in length
* 163769 163868: gap of unknown length
* 163869 170777: contig of 6909 bp in length.

FEATURES

Source

1. 170777
/organism="Rhizolophus ferrumuginum"
/mol_type="genomic DNA"
/db_xref="taxon:59479"
/clone_lib="VMRC7-327G16"
/clone_id="VMRC7"
/note="BAC resource: http://bacpac.chori.org/"

1. 59786
/note="clone overlaps with GenBank Accession Number
AC148387 clone VMRC7-122D13 (center project name gc1)"

1. 17243
/note="assembly_fragment
clone_end:other
vector_side:left"

17344..24932
misc_feature
/note="assembly_fragment"
25033..75704
misc_feature
/note="assembly_fragment"
75805..95337
misc_feature
/note="assembly_fragment"
95438..101416
misc_feature
/note="assembly_fragment"
101517..163768
misc_feature
/note="assembly_fragment"

misc_feature

misc_feature

ORIGIN

Query Match 83.3%; Score 15; DB 2; Length 170777;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3 CCACGGACACACTT 17
DB 79415 CCACGGACACACTT 79429

RESULT 11

AC091307/c

LOCUS

DEFINITION

AC091307

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL


```
/db_xref="GI:21325572"
/translation="MPDNFGDPENTDPPEATSDRAEQAAEAAQAQESSPPGQASE
EISPELEAIINDLSDDPDLDDGSAVESVQQLARTEDLQRTVAEVAAYRRERE
RQGIIDTARAGVVTQLPLDLDDLAEQHDGDLNEGKLSKDLINILGGLKSTRE
IGAFPEIHEAVQDLSOGGVKVLGTVLRKGYLGDVRVIRTMVILIGDEBS"
complement(792..2648)
/gene="Cg12800"
complement(792..2648)
/gene="Cg12800"
complement(792..2648)
/note="PF00012:Hsp70 protein"
/codon_start=1
/translation="Molecular chaperone and 70 kDa heat shock
chaperonin protein dnak"
/protein_id="BAC00194.1"
/db_xref="GI:21325573"
/translation="MGRVAGIDLTGNSVSVSVEGEPVIVANAEGRTPPVAFK
NEVIVGOSAKNOAVTNVDRTRSVKRTISGTMKSVAIIDKNYTSOESARTLTKRD
AEAVYLEDVTDAVITTPAFEDSQRQATKAGQIAGLNTLRVNEPRAALAYGLEKQ
BOEQTLLVEDLGGGTFDVSLLEIGDVEVRATSGDNELGDDWDQRIYDMLVEKQS
ENGIIDTKRMALQRLREAAEKAKIELSSQSANILPYITVDKPNPLFLEDTLSRA
EFORITODLARTKTPFNQVVDGAVSSEIDHVLVVGSTMPAVTELKELTTSRE
PNKGNPDEVAVGAAIQAGVLRGEVDVLLDVTPLSGIETKGMVTLIERNTTI
PTRRSTFTTAEDNOPSVQIOVQGEREITAKKLIGSPELGGIAPAPRGVQIETVF
DIANGIVATYAKDKGTGKENTTTTIDGSLSDSDIDRMIKDAEAHDEKKRREOE
VRNNAESLVYQTKRFVENSEKVSSEDLKAVEBAKVEBALGEDIKAIKAVEKLN
TESQEMGKAIYEADAAAGATQADAGAGADNVVAEVEDDAADGEDKK"
3084..4418
/gene="Cg12801"
3084..4418
/gene="Cg12801"
/note="PF01594:Domain of unknown function DUF20"
/codon_start=1
/translation="Predicted permease"
/product="Predicted permease"
/protein_id="BAC00195.1"
/db_xref="GI:21325574"
/translation="MPDTVLDAIEGNPEPLNTQKSAVPNNAETVEADSEGTTRSV
ILGRDGRMLSGMALRFIVLAGVIALRMKGFIVGILPVALLVCTVILMPVKMLR
EHKIPALAVITITLGFALIGVFAIAPSVSQTQVQVQDAITVGIEDLMDVQGP
FNLDISQFEGALNDLSMLSSQSTISVSGSITSLVITLAVMLVITLFFLLKQG
PRPLPMKRGFTGEMAGCWLTFEVLTRTNLTAGFRALVSMVDALFIGILLIAGP
MALALAVITTFAGFTPIVGAVSAGALVITLALVTNGTNALLVILVLANQOIEGNTL
SPFSQKAMLLAAVLLSVTSGTMEGIIIGAFIAPVAATLAWVRHSEMAVLRAG
EITSDDIEIATTKGSPTSINGOETLAIIRSRFALIGRKKTSATAEKADS"
complement(4462..5097)
/gene="Cg12802"
complement(4462..5097)
/gene="Cg12802"
/EC_number="3.2.2.9"
/EC_number="3.2.2.16"
/codon_start=1
/translation="Nucleoside phosphorylase"
/product="Nucleoside phosphorylase"
/protein_id="BAC00196.1"
/db_xref="GI:21325575"
/translation="MKLFOSCFLLNNSACNLVVLCHVTMTETLTVSATTEEAAYLDP
GIDLVTGIGTATMTILKELATREVLPAIVINIGTALVDELAGVLEIYVLDQD
FSSRLIEMTGKPCSGNSTLATSQHPVASIATNSPFIABSTENHILATRASLDMEG
AALYGVAKHGVPTTLKQVSDSDEBASGMPVADAGARQLAEAKERK"
complement(5094..6296)
/gene="Cg12803"
complement(5094..6296)
/gene="Cg12803"
complement(5094..6296)
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAC00197.1"
/db_xref="GI:21325576"
/translation="MSAPPSWATPKQTPRAOPRAIPVSLGRNNMPLFIDALHRS
KKYFAHSELLIGEFAGLSLPHPTAKAADIDAREFIRQMEGDDVVEVAIRMS
SPVLGKTEVPVRLTINTREILVFAQVEDMSSLHREFSOLSGPTAEVVAKHVLSLR
SLSDNDLSKAVLVVDMFLKHPNSGLKRAVAVEGVHTKMLENHRVLIETTVADKRGEP
```

```
gene
CDS
complement(6202..9537)
/gene="Cg12804"
complement(6202..9537)
/gene="Cg12804"
complement(6202..9537)
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAC00198.1"
/db_xref="GI:21325577"
/translation="MTSEQALPPIHPGQRLRIOLINMGTFHGVTDIVPTREGILVT
GSGSGSKSLTALITAVLLPOGKAPNSAQAONTBPNKRSIVTVIRGMAQOEPLQ
DOIIVSTYLRPRATISLVGLTTSNGSCVETTLVAITLKSGHNLTSISYVCYFPDO
DINALDLFKEDIDRQIAPAAKEAIFESQSHVFGFRFSRIGISSEBALHLHRQS
ADLOSILDFKEDYMLVEPDTPSIKTAVEQODEGAVEQVEDIKRQIHLLDPLVQ
KORREKAOQSKHANAALKALPTVGNRIKKEEQETLVQFTVEQOTQAKSKVESAKLET
DRAREMETLHNNVQIYGAHQIILSAKEGADKRTISTARAGLDALVKGIGGAP
EBAEELIELNNAARLITVDEPPARALSESAGORVENDRTARVDEFKAAOELSLSG
SSNTEYRLQVBEMLQDGVSPRMPAGELIDRNMAWEPVQIIGGPAEMLVPS
HGLPEVRDVAWAKHLAALLKNGVVTTEYKTSRPADSLRKDVSPFRDYNQ
ELGKFNICVATPELSALGPRDQVITLGVKPAQOTGDPPTTWEKDRRLGURS
TYRAGSTNDAKVETLRETVAKGAVVQADNLIANRAELRELEROYQASOELIKYSW
ACIDVESADAALIELDLLEELNPEATELSRBEAKQTLARVAANDVQTLAREL
ASNLKRAETELKRLSELPVAVSEPIAEVVKLPLANRRAVAAVNDQTLAREL
DKQIDANEAELRCENQIVIGLRSYIETTPARADIQAPPEPGEAINLGEIRSRLL
AEFTAKPIELNMEMSTRNLGQISRLRDARREIIEETIPINSLASHDSFNEGFRLND
IRDQSGPIVREEFOQLDAATSGDLSTKQKQIAEIIISKLASHDSADARNA
TVLDTRRHVRVIGLERSDQATVNTVDSAGQGLVFCCLAALRYQLAEBCA
HYPTVATVILDEAFDPAADPAFTROTNNVHSGFPHVLAATPLKLIQTLDGVYSTIV
STYKPNQAQALQGNSSFSRIEK"
complement(9534..10172)
/gene="Cg12805"
complement(9534..10172)
/gene="Cg12805"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAC00199.1"
/db_xref="GI:21325578"
/translation="KNDOLWEGDTGLTFGSRKALVQLKGPVNALLOHVEVRAITT
DQALNAYANLLELVLDNAGVATFRPANGOEVLGNKNTEAMPKVLRTETLSIF
DTIILILKRELMAPPGERVITYDREIIBOULYLVYDERDRDAKAKAPDDAARRLV
DYSLAKKETPERFEVSPALRQIFADVTAGVAAVEYKFKAAHDGNEBOK"
complement(10169..11656)
/gene="Cg12806"
complement(10169..11656)
/gene="Cg12806"
complement(10169..11656)
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAC00200.1"
/db_xref="GI:21325579"
/translation="MTVYSHALGFKRPROSELESLIRSNPPVVLAVVAOYPOQAI
AKPASELYQLSDFRVLKREBGELEKSPDIDYSDWYKSWPFRPRESSQOTGERVES
EBLIAVDSVQORDNPNRSISARISLTOALDTLDESPTAKRIELERERDRIRB
ROJLEAVIAGSEVLYTTOIGDRVADLLAASIPAPFARRHLSIDNRLRQLQOLD
EDSRGVLEIEIFRGVDLIGSDAGRSFNSFFVLLRERSLSLDRWIRVYGRDEAID
LDSKRTGLVRIFRDMEDASFEVNGEWTGLARHYVTTEPAAESRRMQLARDTDS
AAKKAABAGEVTSLNHMDPLVIRIGMDVRSIAGLKIKNGEERVEDLPEVEBOELDT
EVLMEQIRASEIDPELEBAVSLVLABQSHATTBVLNHPATQGLASIVGLLYLMAR
DGVPTGAAQIVEWESDATHRRIRITGQPIRGINSBDLASEBMDK"
complement(11995..13032)
/gene="Cg12807"
complement(11995..13032)
/gene="Cg12807"
complement(11995..13032)
/EC_number="1.1.1.1"
/note="PF00107:zinc-binding dehydrogenases
TIGR00692:cdh: L-threonine 3-dehydrogenase"
/codon_start=1
/translation="1
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 128.44 Seconds
(without alignments)
5630.828 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19

Sequence: 1 tctagcgcgcacacgcgtcg 19

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues
Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	1489	9	AG109756 Pan trogl
2	17	89.5	163	6	CD323908 StrPu537.
3	17	89.5	333	6	CD324115 StrPu537.
4	17	89.5	461	6	CD339324 StrPu536.
5	17	89.5	570	6	CD912036 G550.113A
6	17	89.5	676	6	BM607198 170006870
7	17	89.5	752	4	BI144572 602909436
8	17	89.5	942	7	CF546859 AGERNCOURT
9	16	84.2	298	1	AI537835 tp26a08.x
10	16	84.2	527	8	AQ644580 HS 5104 A
11	16	84.2	546	9	CE546398 tigt-gss-
12	16	84.2	579	4	BF983915 602307473
13	16	84.2	651	9	AG154721 Pan trogl
14	16	84.2	1091	4	BF983743 602307337
15	16	84.2	251	7	CF842669 PSBH021XB
16	15	78.9	268	6	CD613609 56029053J
17	15	78.9	278	2	AM428561 66804 MAR
18	15	78.9	281	6	CD613605 56028993J
19	15	78.9	410	6	CD613589 56028923J
20	15	78.9	421	2	BF117366 u236b06.Y
21	15	78.9	465	4	BI078010 602874382
22	15	78.9	482	5	BO753281 EBAN01 SO
23	15	78.9	484	4	BI339263 364169 MA
24	15	78.9	508	2	BE227626 894033G09

25	15	78.9	524	6	CB066279	PVB12B05
26	15	78.9	544	1	AJ447864	AJ447864
27	15	78.9	587	2	BF528035	BF528035
28	15	78.9	587	6	CB427482	CB427482
29	15	78.9	605	4	BG712523	BG712523
30	15	78.9	607	5	BQ038392	BQ038392
31	15	78.9	627	7	CF571178	CF571178
32	15	78.9	638	4	BI066464	BI066464
33	15	78.9	638	7	CN274909	CN274909
34	15	78.9	639	4	BI851715	BI851715
35	15	78.9	680	9	BX905787	BX905787
36	15	78.9	680	9	LBFA021A05	LBFA021A05
37	15	78.9	682	1	AJ452357	AJ452357
38	15	78.9	695	4	BG764608	BG764608
39	15	78.9	715	6	BY707912	BY707912
40	15	78.9	723	6	CB457166	CB457166
41	15	78.9	730	5	BU251290	BU251290
42	15	78.9	736	6	CD342162	CD342162
43	15	78.9	758	4	BI156664	BI156664
44	15	78.9	762	7	CO879708	CO879708
45	15	78.9	771	4	BI660197	BI660197

ALIGNMENTS

RESULT 1
AG109756
LOCUS
Definition
Accession
Version
Keywords
Source
Organism
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE
ORIGIN
Query Match
Best Local Similarity

AG109756 1489 bp DNA linear GSS 03-NOV-2001
Pan troglodytes DNA, clone: PTB-115E10.R, genomic survey sequence.
AG109756
AG109756.1 GI:16730275
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of library PTB
2 (bases 1 to 1489)
Unpublished
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan
E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170
Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: M13rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 1489
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-115E10.R"
/sex="male"
/cell_type="lymphoblast"
/clone_id="PTB Chimpanzee Male BAC Library"

100.0%; Score 19; DB 9; Length 1489;
100.0%; Pred. No. 0.27;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGT 19
|||||
|||||

Db 739 TGTAGCGGCCACGCGT 757

RESULT 2
LOCUS CD323908
DEFINITION CD323908 163 bp mRNA linear EST 17-SEP-2003
StrP537.002270 Sea urchin embryo 20hr blastula stage cDNA library
MPMG537 Strongylocentrotus purpuratus cDNA clone
CD323908
CD323908.1 GI:34795969
EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 163)
Pouska,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.
Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters
Genome Res. 13 (12), 2736-2746 (2003)
Contact: Pouska AJ
laboraty 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inmestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: pouska@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (<http://www.rzpd.de>)
PCR Primers
FORWARD: 5' CCCAGGCTTACACTTATGCTTCCGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCGACCTGCGAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5' GCTATTACGCGACCTGCGAAGGGGATGTG 3' (M13FSP)
High quality sequence stop: 163.
Location/Qualifiers
1. 163
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="CALIP537E1217;MPI_537_17E12"
/issue_type="whole embryo"
/dev_stage="embryonic 20hr"
/lab_host="E.coli, XLI blue"
/clone_lib="Sea urchin embryo 20hr blastula stage cDNA
library MPMG537"
/note="Vector: pSport1; site 1: NotI; site 2: SalI; Random
primed and directionally cloned in pSport1 vector using a
NotI (5'-TGACCTTCTAGATCGGACGCGCC (T)15-3' and a
SalI 5'-TCGACCCACGCGTCG-3' adapters (Gibco BRL) "

ORIGIN
Query Match 89.5%; Score 17; DB 6; Length 163;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGT 17
|||||
|||||

Db 135 TGTAGCGGCCACGCGT 151
|||||
|||||

RESULT 3
LOCUS CD324115/c
DEFINITION CD324115 333 bp mRNA linear EST 17-SEP-2003
StrP537.002477 Sea urchin embryo 20hr blastula stage cDNA library
MPMG537 Strongylocentrotus purpuratus cDNA clone
CALIP537K1047;MPI_537_47K10 3', mRNA sequence.
CD324115
CD324115.1 GI:34796176
EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 333)
Pouska,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.
Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters
Genome Res. 13 (12), 2736-2746 (2003)
Contact: Pouska AJ
laboraty 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inmestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: pouska@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (<http://www.rzpd.de>)
PCR Primers
FORWARD: 5' CCCAGGCTTACACTTATGCTTCCGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCGACCTGCGAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5' GCTATTACGCGACCTGCGAAGGGGATGTG 3' (M13FSP)
High quality sequence stop: 333.
Location/Qualifiers
1. 333
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="CALIP537K1047;MPI_537_47K10"
/issue_type="whole embryo"
/dev_stage="embryonic 20hr"
/lab_host="E.coli, XLI blue"
/clone_lib="Sea urchin embryo 20hr blastula stage cDNA
library MPMG537"
/note="Vector: pSport1; site 1: NotI; site 2: SalI; Random
primed and directionally cloned in pSport1 vector using a
NotI (5'-TGACCTTCTAGATCGGACGCGCC (T)15-3' and a
SalI 5'-TCGACCCACGCGTCG-3' adapters (Gibco BRL) "

ORIGIN
Query Match 89.5%; Score 17; DB 6; Length 333;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGT 17
|||||
|||||

Db 30 TGTAGCGGCCACGCGT 14
|||||
|||||

RESULT 4
LOCUS CD339324 461 bp mRNA linear EST 17-SEP-2003
DEFINITION ScP536.003145 Sea urchin embryo 40hr gastrula stage cDNA library
MPM6536 Strongylocentrotus purpuratus cDNA clone
CALTP536H0915.MPI_536_15H9 3', mRNA sequence.
ACCESSION CD339324
VERSION CD339324
KEYWORDS Strongylocentrotus purpuratus
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
REFERENCE Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.
AUTHORS 1 (bases 1 to 461)
Pouska,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A., Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.
TITLE Generation, annotation, evolutionary analysis, and database integration of 20,000 unique sea urchin EST clusters
JOURNAL Genome Res. 13 (12), 2736-2746 (2003)
COMMENT Contact: Pouska AJ
laboraty 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Imestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: pouska@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting (ONP) to reduce sequencing redundancy. According to the ONP procedure, clones that display the same hybridisation matrix with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per ONP cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of which the above EST is generated is available at the sea urchin project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA clones and filters are distributed via the Resource Center/Primary Database of the German Human Genome Project (<http://www.rzpd.de>)
PCR Primers
FORWARD: 5' CCCAGGCTTACCTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCCAGCTGCGCAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5' GCTATTACGCCAGCTGCGCAAGGGGATGTG 3' (M13FSP)
High quality sequence stop: 461.
Location/Qualifiers
1..461
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="CALTP536H0915.MPI_536_15H9"
/tissue_type="whole embryo"
/dev_stage="embryonic 40hr"
/lab_host="E.coli, XLI blue"
/clone_lib="Sea urchin embryo 40hr gastrula stage cDNA library MPM6536"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; Random primed and directionally cloned in pSPORT1 vector using a NotI (5'-GCACTAGTCTTAGATCGGAGGAGGCGGCC (T)15-3' and a SalI 5'-TCGACCAACGCTCCG-3'adapters (Gibco BRL)."
ORIGIN
Query Match 89.5%; Score 17; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS CD912036 570 bp mRNA linear EST 14-JUL-2003
DEFINITION G550.113A12P010522 G550 Triticum aestivum cDNA clone G550113A12, mRNA sequence.
ACCESSION CD912036
VERSION CD912036.1 GI:32686360
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
AUTHORS 1 (bases 1 to 570)
Genoplante.
TITLE Unpublished (2003)
JOURNAL Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers
1..570
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="rectal"
/db_xref="taxon:4565"
/clone="G550113A12"
/tissue_type="grain (550 degrees per day after pollination)"
/clone_lib="G550"
ORIGIN
Query Match 89.5%; Score 17; DB 6; Length 570;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
LOCUS BM607198/c 676 bp mRNA linear EST 25-FEB-2002
DEFINITION 17000687084830 A.Gam.ad.cDNA.bloddi Anopheles gambiae cDNA clone 19600449683177 5', mRNA sequence.
ACCESSION BM607198
VERSION BM607198.1 GI:18905302
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE 1 (bases 1 to 676)
Holt,R.A., Jin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: holtra@celera.com
Plate: N001004A2S row: L column: 11
Seq primer: M13 Reverse.
Location/Qualifiers
1..676
/organism="Anopheles gambiae"

/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="1960049683177"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not I site 18 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"

ORIGIN

Query Match 89.5%; Score 17; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCCGCGT 17
|||||
Db 56 TGTAGCGGCCCGCCGCGT 40

RESULT 7 752 bp mRNA linear EST 05-JUL-2001
B1144572/c 602909436F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5050424 5',
LOCUS mRNA sequence.
DEFINITION B1144572
ACCESSION B1144572
VERSION B1144572.1 GI:14604573
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 752)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Cloned through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLAM1137 row: h column: 09
High quality sequence stop: 445.

FEATURES
Location/Qualifiers
1..752
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5050424"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: liver; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 89.5%; Score 17; DB 4; Length 752;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCCGCGT 17
|||||

Db 19 TGTAGCGGCCCGCCGCGT 3

RESULT 8 942 bp mRNA linear EST 22-SEP-2003
CF546859/c A6ENECOURT_15607895 NICHG_XGC_Kid1 Xenopus laevis cDNA clone
LOCUS IMAGE:7008145 5', mRNA sequence.
DEFINITION CF546859
ACCESSION CF546859.1 GI:34883691
VERSION EST.
KEYWORDS Xenopus laevis (African clawed frog)
SOURCE Xenopus laevis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 942)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Cloned through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLAM14711 row: g column: 23
High quality sequence stop: 642.

FEATURES
Location/Qualifiers
1..942
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:7008145"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: PCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.2 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."

ORIGIN

Query Match 89.5%; Score 17; DB 7; Length 942;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCCGCGT 17
|||||
Db 32 TGTAGCGGCCCGCCGCGT 16

RESULT 9 298 bp mRNA linear EST 13-MAY-1999
A1537835/c tp26a08.x1 NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE:2188886 3',
LOCUS similar to TR:Q13488 Q13488 SPECIFIC 116-KDA VACUOLAR PROTON PUMP
DEFINITION SUBMIT. :, mRNA sequence.
ACCESSION A1537835
VERSION A1537835.1 GI:4451970
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 298)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.illn.gov/bdtp/image/image.html
 Insert length: 1961 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 168
 POLY-A-NO.

FEATURES

source

1. .298
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2188886"
 /tissue_type="poorly differentiated adenocarcinoma with
 signed ring cell features"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Gas4"
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCCGCGCT 17
 |||||
 244 GTAGCGGCCCGCGCT 229

RESULT 10 A0464580 527 bp DNA linear GSS 23-APR-1999
LOCUS A0464580/c
DEFINITION HS_5104_A1_G12_T7A RPCT-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=680 Col=23 Row=M, genomic survey sequence.

ACCESSION A0464580
VERSION A0464580.1 GI:4641675
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 527)
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCT-11. For BAC
 library availability, please contact Pieter de Jong

(pietere@dejong.med.bufileo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.bufileo.edu/ordering_bac.htm)
 or from Research Genetics (<http://www.htec.washington.edu>)
info@resgen.com. BAC end Web Server:
<http://www.htec.washington.edu>
 Plate: 680 row: M column: 23
 Seq primer: 17
 Class: BAC ends
 High quality sequence stop: 527.
 Location/Qualifiers

FEATURES

source

1. .527
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=680 Col=23 Row=M"
 /sex="male"
 /clone_lib="RPCT-11 Human Male BAC Library"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"

ORIGIN

Query Match 84.2%; Score 16; DB 8; Length 527;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCGCGCTGG 19
 |||||
 24 AGCGGCCCGCGCTGG 9

RESULT 11 CE548398 546 bp DNA linear GSS 28-SEP-2003
LOCUS CE548398/c
DEFINITION tigr-gss-dog-17000327426188 Dog Library Canis familiaris genomic,
 genomic survey sequence.
ACCESSION CE548398
VERSION CE548398.1 GI:36665179
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
AUTHORS 1 (bases 1 to 546)
 Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirkness@tigr.org
 Class: shotgun.

FEATURES

source

1. .546
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /note="Site 1: BstXI; Libraries were prepared from
 peripheral blood"

ORIGIN

Query Match 84.2%; Score 16; DB 9; Length 546;
Best Local Similarity 100.0%; Pred. No. 20;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCGACGCGTGG 19
|||||

Db 268 AGCGGCCCGACGCGTGG 253

RESULT 12
BF983915/c 579 bp mRNA linear EST 23-JAN-2001
LOCUS 602307473F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4398619 5',
DEFINITION mRNA sequence.

ACCESSION BF983915
VERSION BF983915.1 GI:12386727
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 579)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM10101 row: a column: 20
High quality sequence stop: 579.

FEATURES
source
1. 579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4398619"
/tissue_type="nodular adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site 1: NCI; Site 2: Sali; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH-MGC Library."

ORIGIN

Query Match 84.2%; Score 16; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCCGACGCGT 17
|||||

Db 17 GTAGCGGCCCGACGCGT 2

RESULT 13
AG154721/c 651 bp DNA linear GSS 09-JAN-2002
LOCUS AG154721
DEFINITION Pan troglodytes DNA, clone: RP43-018N05.T7, genomic survey
sequence.
ACCESSION AG154721
VERSION AG154721.1 GI:16684399
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
1

REFERENCE
AUTHORS

TITLE Totoki, Y., Watanabe, H. and Sakaki, Y.
JOURNAL BAC end sequences of library RPCI-43
REFERENCE 2 (bases 1 to 651)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@cc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS

Sequencing: T7
LIBRARY Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. 651
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-018N05.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lhb="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 84.2%; Score 16; DB 9; Length 651;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCGACGCGTGG 19
|||||

Db 500 AGCGGCCCGACGCGTGG 485

RESULT 14
BF983743/c 1091 bp mRNA linear EST 23-JAN-2001
LOCUS 60230733F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4398600 5',
DEFINITION mRNA sequence.
ACCESSION BF983743
VERSION BF983743.1 GI:12386555
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 1091)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM10101 row: i column: 09
High quality sequence stop: 689.

FEATURES

source
1. 1091
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

Search completed: February 10, 2005, 17:01:53
 Job time : 132.44 secs

ORIGIN

/clone="IMAGE:4398800"
 /issue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_88"
 /note="Organ: small_intestine; Vector: PCMV-SPORT6;
 Site_1: Not; Site_2: SalI; Cloned unidirectionally;
 oligo-dt primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."

Query Match 84.2%; Score 16; DB 4; Length 1091;
 Best Local Similarity 100.0%; Pred.No.19;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGCCACGCGT 17
 |||||
 DB 17 GTAGCGCCACGCGT 2

RESULT 15 251 bp mRNA linear EST 30-OCT-2003
 CF842669 pSHB021B20f USDA-IFARS:Expression of Phytophthora sojae genes
 LOCUS during infection and propagation_sHB Phytophthora sojae cDNA clone
 DEFINITION SHB021B20 5, mRNA sequence.

ACCESSION CF842669 GI:38058323
 VERSION CF842669
 KEYWORDS EST.

SOURCE Phytophthora sojae
 ORGANISM Phytophthora sojae
 Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 Phytophthora.

REFERENCE 1 (bases 1 to 251)

AUTHORS Tyler,B. Not Published
 TITLE Unpublished (2003)
 JOURNAL
 COMMENT Contact: Tyler B

TYLER lab

VBI 1880 Pratt Dr., Blackebury, VA 24061, USA

Tel: 540-231-7318

Email: bmtyley@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 021 row: B column: 20

Seq primer: BK reverse primer

High quality sequence stop: 251.

location/Qualifiers

1. .251

/organism="Phytophthora sojae"

/mol_type="mRNA"

/db_xref="taxon:67593"

/clone="SHB021B20"

/issue_type="mycelium"

/cell_line="P6497"

/dev_stage="48 hr. post infection stage"

/lab_host="Soybean plant"

/clone_lib="USDA-IFARS:Expression of Phytophthora sojae
 genes during infection and propagation_sHB"

/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 78.9%; Score 15; DB 7; Length 251;

Best Local Similarity 100.0%; Pred.No.91;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GCGGCCACGCGTGG 19
 |||||
 DB 62 GCGGCCACGCGTGG 76

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 5.8365 Seconds
(without alignments)
5324.730 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19
Sequence: 1 ttttagcgccacgcgctgg 19

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	19	4	US-09-872-051-9	Sequence 9, Appl1
2	100.0	498	4	US-09-872-051-7	Sequence 7, Appl1
3	78.9	843	4	US-09-902-540-9093	Sequence 9093, Ap
4	78.9	1405	1	US-08-390-162-3	Sequence 3, Appl1
5	78.9	1405	1	US-08-685-945B-3	Sequence 3, Appl1
6	78.9	2107	1	US-08-390-162-1	Sequence 1, Appl1
7	78.9	2107	1	US-08-685-945B-1	Sequence 1, Appl1
8	78.9	2558	4	US-09-902-540-6342	Sequence 6342, Ap
9	78.9	2560	4	US-09-902-540-400	Sequence 400, App
10	78.9	10915	4	US-09-902-540-972	Sequence 972, App
11	78.9	1083	4	US-09-902-540-5440	Sequence 5440, Ap
12	73.7	1092	4	US-09-902-540-8179	Sequence 8179, Ap
13	73.7	1209	4	US-09-826-509-556	Sequence 556, App
14	73.7	1376	4	US-09-016-434-1361	Sequence 1361, Ap
15	73.7	1394	2	US-08-068-729-3	Sequence 3, Appl1
16	73.7	1394	3	US-09-255-671-3	Sequence 3, Appl1
17	73.7	1394	3	US-09-395-366-3	Sequence 3, Appl1
18	73.7	1488	3	US-09-096-776B-3	Sequence 3, Appl1
19	73.7	1488	4	US-09-923-922-3	Sequence 3, Appl1
20	73.7	2457	4	US-09-949-016-1330	Sequence 1330, Ap
21	73.7	2457	4	US-09-949-016-1331	Sequence 1331, Ap
22	73.7	2534	3	US-09-096-776B-1	Sequence 1, Appl1
23	73.7	2534	3	US-09-923-922-1	Sequence 1, Appl1
24	73.7	2640	4	US-08-684-932A-37	Sequence 37, Appl1
25	73.7	2640	4	US-09-618-304B-1	Sequence 1, Appl1
26	73.7	2654	4	US-09-949-016-2404	Sequence 2404, Ap
27	73.7	2654	4	US-09-949-016-2405	Sequence 2405, Ap

C 28	14	73.7	2655	4	US-09-016-434-1094	Sequence 1094, Ap
C 29	14	73.7	2655	4	US-09-023-655-916	Sequence 916, App
C 30	14	73.7	2706	4	US-09-949-016-2975	Sequence 2975, Ap
C 31	14	73.7	2706	4	US-09-949-016-2976	Sequence 2976, Ap
C 32	14	73.7	3465	4	US-09-489-039A-553	Sequence 553, App
C 33	14	73.7	3603	4	US-09-902-540-3266	Sequence 3266, App
C 34	14	73.7	11917	4	US-09-949-016-13073	Sequence 13073, A
C 35	14	73.7	11917	4	US-09-949-016-13072	Sequence 13072, A
C 36	14	73.7	13256	4	US-09-902-540-1006	Sequence 1006, Ap
C 37	14	73.7	14638	4	US-09-902-540-1106	Sequence 1106, Ap
C 38	14	73.7	15061	4	US-09-949-016-14717	Sequence 14717, A
C 39	14	73.7	15061	4	US-09-949-016-14718	Sequence 14718, A
C 40	14	73.7	15859	4	US-09-949-016-14146	Sequence 14146, A
C 41	14	73.7	15859	4	US-09-949-016-14147	Sequence 14147, A
C 42	14	73.7	19954	4	US-09-902-540-1150	Sequence 1150, Ap
C 43	14	73.7	36800	4	US-08-311-733A-139	Sequence 139, App
C 44	14	73.7	41310	4	US-09-902-540-1284	Sequence 1284, Ap
C 45	13	68.4	36	3	US-08-434-099A-36	Sequence 36, Appl1

ALIGNMENTS

RESULT 1
US-09-872-051-9
; Sequence 9, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Det
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(19)
; OTHER INFORMATION: Zea maize genomic and vector DNA
US-09-872-051-9
Query Match 100.0%; Score 19; DB 4; Length 19;
Best local Similarity 100.0%; Pred. No. 0.16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
C 1 ttttagcgccacgcgctgg 19
Db 1 ttttagcgccacgcgctgg 19
RESULT 2
US-09-872-051-7
; Sequence 7, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine

APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMG32(hk603) and Composition and Methods for Detecting
TITLE OF INVENTION: Thereof
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 498
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)-(498)
OTHER INFORMATION: 1-304 Zea maize genomic DNA
OTHER INFORMATION: 305-349 construct vector DNA
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA
US-09-872-051-7

Query Match 100.0%; Score 19; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.092; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGTGG 19
DB 296 TGTAGCGGCCACGCGTGG 314

RESULT 3
US-09-902-540-9093/C
Sequence 9093, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 9093
LENGTH: 843
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-9093

Query Match 78.9%; Score 15; DB 4; Length 843;
Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCACGCGC 16
DB 693 GTAGCGGCCACGCGC 679

RESULT 4
US-08-390-162-3/C
Sequence 3, Application US/08390162
Patent No. 5576192
GENERAL INFORMATION:
APPLICANT: Ichikawa, Atsushi

APPLICANT: Natumiya, Shuh
TITLE OF INVENTION: Proctoglandin E Receptors, Their DNA and
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,162
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/024179
FILING DATE: 23-FEB-1993
APPLICATION NUMBER: JP 036580-1992
FILING DATE: 24-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 064889-1992
FILING DATE: 23-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: FortiB, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 04221-0020-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1405 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-390-162-3

Query Match 78.9%; Score 15; DB 1; Length 1405;
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGC 15
DB 631 TGTAGCGGCCACGCGC 617

RESULT 5
US-08-685-945B-3/C
Sequence 3, Application US/08685945B
Patent No. 5804415
GENERAL INFORMATION:
APPLICANT: Ichikawa, Atsushi
APPLICANT: Natumiya, Shuh
TITLE OF INVENTION: Proctoglandin E Receptors, Their DNA and
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,945B
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/024179
FILING DATE: 23-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 035580-1992
FILING DATE: 24-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 064889-1992
FILING DATE: 23-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 04221-0020-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1405 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-685-945B-3

Query Match 78.9%; Score 15; DB 1; Length 1405;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGC 15
Db 631 TGTAGCGGCCACGC 617

RESULT 6
US-08-390-162-1/c
Sequence 1, Application US/08390162
Patent No. 5576192
GENERAL INFORMATION:
APPLICANT: Ichikawa, Atsushi
APPLICANT: Natumiya, Shuh
TITLE OF INVENTION: Prostoglandin B Receptors, Their DNA and
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,162
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/024179
FILING DATE: 23-FEB-1993
APPLICATION NUMBER: JP 036580-1992
FILING DATE: 24-FEB-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 064889-1992
FILING DATE: 23-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 04221-0020-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2107 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-390-162-1

Query Match 78.9%; Score 15; DB 1; Length 2107;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGC 15
Db 644 TGTAGCGGCCACGC 630

RESULT 7
US-08-685-945B-1/c
Sequence 1, Application US/08685945B
Patent No. 5804415
GENERAL INFORMATION:
APPLICANT: Ichikawa, Atsushi
APPLICANT: Natumiya, Shuh
TITLE OF INVENTION: Prostoglandin B Receptors, Their DNA and
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,945B
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/024179
FILING DATE: 23-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 036580-1992
FILING DATE: 24-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 064889-1992
FILING DATE: 23-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 04221-0020-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2107 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-685-945B-1

Query Match 78.9%; Score 15; DB 1; Length 2107;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCG 15
Db 644 TGTAGCGGCCACGCG 630

RESULT 8
US-09-902-540-6342/C
Sequence 6342, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 6342
LENGTH: 2558
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(2558)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-6342

Query Match 78.9%; Score 15; DB 4; Length 2558;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTAGCGGCCACGCG 16
Db 162 GTAGCGGCCACGCG 148

RESULT 9
US-09-902-540-400/C
Sequence 400, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 400
LENGTH: 2560
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(2560)

OTHER INFORMATION: unsure at all n locations
US-09-902-540-400

Query Match 78.9%; Score 15; DB 4; Length 2560;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTAGCGGCCACGCG 16
Db 164 GTAGCGGCCACGCG 150

RESULT 10
US-09-902-540-972
Sequence 972, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 972
LENGTH: 10915
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-972

Query Match 78.9%; Score 15; DB 4; Length 10915;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTAGCGGCCACGCG 16
Db 152 GTAGCGGCCACGCG 166

RESULT 11
US-09-902-540-5440/C
Sequence 5440, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 5440
LENGTH: 1083
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-5440

Query Match 73.7%; Score 14; DB 4; Length 1083;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TAGCGGCCACGCG 16
Db 355 TAGCGGCCACGCG 342

RESULT 12
US-09-902-540-8179/c
; Sequence 8179, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8179
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8179

Query Match 73.7%; Score 14; DB 4; Length 1092;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCACGCG 15
Db 921 GTAGCGGCCACGCG 908

RESULT 13
US-09-826-509-556/c
; Sequence 556, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Brunisma, Karin
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/1170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 556
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-826-509-556

Query Match 73.7%; Score 14; DB 4; Length 1209;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCACGCG 16
Db 536 TAGCGGCCACGCG 523

RESULT 14
US-09-016-434-1361/c
; Sequence 1361, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1361:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g410208
US-09-016-434-1361

Query Match 73.7%; Score 14; DB 4; Length 1376;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCACGCG 16
Db 610 TAGCGGCCACGCG 597

RESULT 15
US-08-068-729-3/c
; Sequence 3, Application US/08068729
; Patent No. 5985597
; GENERAL INFORMATION:
; APPLICANT: Ford-Hutchinson, Anthony
; APPLICANT: Funk, Colin
; APPLICANT: Grygorczyk, Richard
; APPLICANT: Metters, Kathleen
; TITLE OF INVENTION: DNA Encoding Prostaglandin Receptor EPI
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOHN W. WALLIN III
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/068,729
/ FILING DATE: 26-MAY-1993
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: WALLEN, JOHN W III
/ REGISTRATION NUMBER: 35,403
/ REFERENCE/DOCKET NUMBER: 19012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 594-3905
/ TELEFAX: (908) 594-4720
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1394 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-08-068-729-3

Query Match 73.7%; Score 14; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCCAAGCG 16
|||
Db 610 TAGCGGCCCAAGCG 597

Search completed: February 10, 2005, 08:49:48
Job time: 6.83865 secs